

FIGURE 1

CCAGGTCCAAC TGACCTCGGTTCTATCGATTGAATCCCCGGGGATCCTCTAGAGATCCCT
CGACCTCGACCCACGCGTCCGCCAAGCTGGCCCTGCACGGCTGCAAGGGAGGCTCCTGTGGA
CAGGCCAGGCAGGTGGGCTCAGGAGGTGCTCCAGGCGCCAGTGGGCTGAGGCCCCAGC
AAGGGCTAGGCTCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGGCCACGCTTGGC
TCCAGCAGCATCAGCAGCCCCCAGGACCGGGGGAGGCACAGGTGGCCCCCACCACCCGGAGG
AGCAGCTCCTGCCCCCTGTCCGGGGATGACTGATTCTCCTCCGCCAGGCCACCCAGAGGAGA
AGGCCACCCCGCTGGAGGCACAGGCCATGAGGGGCTCTCAGGAGGTGCTGCTGATGTGGCT
TCTGGTGTGGCAGTGGGCGGCACAGAGCAGCCTACCGGCCCGGCCGTTAGGGTGTGTGCT
GTCCCGGGCTCACGGGACCCCTGTCTCCAGTCTGTTCTGTCAGCGTGTGTACCAGCCCTTC
TCACCACCTGCGACGGGCACCGGCCCTGCAGCACCTACCGAACCATTTATAGGACCGCCTAC
CGCCGACGCCCTGGGCTGGCCCCCTGCCAGGCCTCGCTACGCGTGTGCCCCGGCTGGAAGAG
GACCAGCGGGCTTCTTGGGGCTGTGGAGCAGCAATATGCCAGCCGCATGCCGGAACGGAG
GGAGCTGTGTCCAGCCTGGCCGTGCGCGCTGCCCTGCAGGATGGCGGGGTGACACTTGCCAG
TCAGATGTGGATGAATGCAGTGTAGGAGGGGCGGCTGTCCCCAGCGCTGCATCAACACCGC
CGGCACTTACTGGTGCCAGTGTGGGAGGGGCACAGCCTGTCTGCAGACGGTACACTCTGTG
TGCCCAAGGGAGGGCCCCCAGGGTGGCCCCAACCCGACAGGAGTGGACAGTGAATGAAG
GAAGAAGTGCAGAGGCTGCAGTCCAGGGTGGACCTGTGAGGAGAAGCTGCAGCTGTGTGCT
GGCCCCACTGCACAGCCTGGCCTCGCAGGCACTGGAGCATGGGCTCCCGGACCCCGGCAGCC
TCCTGGTGCACTCCTTCAGCAGCTCGGCCGCATCGACTCCCTGAGCGAGCAGATTTCCTTC
CTGGAGGAGCAGCTGGGGTCTGCTCTGCAAGAAAGACTCGTGA CTGCCACAGCGCCCCAGG
CTGGACTGAGCCCTCACGCCGCCCTGCAGCCCCCATGCCCTTGCCCAACATGCTGGGGGTC
CAGAAGCCACCTCGGGGTGACTGAGCGGAAGGCCAGGCAGGCGCTTCCTCCTTTCTCCTC
CCCTTCCCTCGGGAGGGTCCCCAGACCCTGGCATGGGATGGGCTGGGATTTTTTTTGTGAAT
CCACCCCTGGCTACCCCCACCCCTGGTTACCCCAACGGCATCCCAAGGCCAGGTGGGCCCTCA
GCTGAGGGAAGGTACGAGTTCCTCTGCTGGAGCCTGGGACCCATGGCACAGGCCAGGCAGCC
CGGAGGCTGGGTGGGCTCAGTGGGGCTGTGCTGCTGACCCCCAGCACATAAAATGAAA
CGTGA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGGCCGCACTCT
AGAGTCGACCTGCAGAAGCTTGGCCGCATGGCCCACTTGTTTATTGCAGCTTATAATGGT
TACAAAT

FIGURE 2

MTDSPPPGHPEEKATPPGGTGHEGLSGGAADVASGVGSGRHRARLPARPLGCVLSSRAHGDPV
SESFVQRVYQPFLLTTCDGHRACSTYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGAC
GAaicQPPCRNGGSCVQPGRCRCPAGWRGDTQSDVDECSARRGGCPQRCINTAGSYWCQCW
EGHSLSADGTL CVPKGGPPRVAPNPTGVDSAMKEEVQRLQSRVDLLEEKQLVLAPLHSLAS
QALEHGLPDPGSLLVHSFQQLGRIDSLSEQISFLEEQLGSCSCKKDS

Signal sequence:

amino acids 1-19

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 93-97, 270-274

N-myristoylation sites.

amino acids 19-25, 78-84, 97-103, 100-106, 103-109, 157-163,
191-197, 265-271

Amidation site.

amino acids 26-30

Aspartic acid and asparagine hydroxylation site.

amino acids 152-164

Cell attachment sequence.

amino acids 130-133

EGF-like domain cysteine pattern signature.

amino acids 123-135

FIGURE 3

CGCTCGCCCCGTCGCCCCCTCGCCTCCCCGCAGAGTCCCCTCGCGGCAGCAGATGTGTGTGGG
GTGAGCCACGGCGGGGACTATGGTGAAATTCGCGGCGCTCACGCACTACTGGCCCCGTGATC
CGGTTCTTGGTGCCCTGGGCATCACCAACATAGCCATCGACTTCGGGGAGCAGGCCCTTGAA
CCGGGGCATTGTCTGTCTCAAGGAGGATGCAGTGCAGATGCTGGCCAGTCCAGGCTGGCGGT
ACTCCCTCATGAAGTCTTTCACGGGTCCCATGAGTGACTTCAAAAATGTGGGCCCTGGTGT
GTGAACAGCAGAGAGACAGGACCAGGCCGCTCTGTGTATGGTGGTGGCAGGGGCCATCGC
TGGCGTCTTTCACACATGTGATGCTTATAGTGATTAGGATACATCATTAATCAAACTGC
ACCATGTGGACGAGTCGGTGGGGAGCAAGACGAGAAGGGCCTTCTGTACTCTCGCCGCTTT
CCTTTCATGGACCAATGGCATGGACCCATGCTGGCATTCTCTTAAACACAAATACAGTTT
CCTGGTGGGATGTGCCCTCAATCTCAGATGTCTAGCTCAGGTTGTTTTGTAGCCATTTTGC
TTCACAGTCACCTGGAATGCCGGGAGCCCCGTGCTCATCCCGATCCTCTCCTGTATCATGGG
GCCTTGTGCGCTGCACCAACCTGTGCTGGGCTACTACAAGAACATTACGACATCATCCC
TGACAGAAGTGGCCCCGAGCTGGGGGGAGATGCAACAATAAGAAAGATGCTGAGCTTCTGGT
GGCCTTTGGCTCTAATTTGCGCCACACAGAGAATCAGTCGGCCATTTGTCAAACCTCTTGGT
TCCCGGGAGCTTGGTGGCAGTTCTGCAGCCACAGAGGAGTCGGCATTTTGACAGCCACATA
CCCTGTGGGTACATAGCCATACGGCTGGTTGACGGAATTCGGTGTGTATCCTGCTTTG
ACAAGAATAACCCGACCAAACTGGTGAGCAGCAGCAACAGTCACGGCAGGCCCATC
AAGAAGTTACCTTCGTCTGCATGGCTCTGTCACTCAGCCTCTGTTTCGTGATGTTTTGGAC
ACCCAAAGTGTGTGAGAAAATCTTGATAGACATCATCGGAGTGGACTTGGCTTTGACAGAAC
TCTGTGTGTTCTTCTTGGCGATCTTCTCCTTCTCCAGTTCAGTCAAGTGGGGCGCAT
CTCACCGGGTGGCTGATGACACTGAAGAAAACCTTCGTCTTGGCCCCAGCTCTGTGCTGG
GATCATCGTCTCTATCGCCAGCCTCGTGGTCTTACCCTACCTGGGGGTGCACGGTGCAGCC
TGGCGTGGGCTCCCTCGCTGGCGGGCTTTGTGGGAGAACTCCACATGGTCCGCATCTGGTGC
TGCTATGTCCTACCGGAAGCAGAAAAAGAAATGGAGAATGAGTCGGCCACGGAGGGGGAAG
CTCTGCCATGACAGACATGCCTCCGACAGAGGAGGTGACAGACATCGTGGAAATGAGAGAG
AGAATGAATAAGGCACGGGACGCCATGGGCACTGCAGGACGGTCAGTCAGGATGACATTC
GGCATCATCTCTCCCTCTCCCATCGTATTTTGTTCCTTTTTTTTGTGTTTTTGTGTAAT
GAAAGAGGCCCTTGATTTAAAGGTTTCGTGTCAATTCTCTAGCATACTGGGTATGCTCACAT
GACGGGGGGACCTAGTGAAATGGTCTTTACTGTTGCTATGTAAAAACAAACGAAACAACTGAC
TTCATACCCCTGCCTCAGGAAAACCCAAAGACACAGCTGCCTCAGCGTTGACGTTGTGTCC
TCCTCCCCTGGACAATCTCCTCTTGGAACCAAAGGAGTGCAGCTGTGCCATCGCGCCTCGGT
CACCCCTGCACAGCAGGCCACAGACTCTCCTGTCCCCCTTCATCGCTCTTAAGAAATCAACAGG
TTAAACCTCGGCTTCCCTTGATTGCTTCCAGTTCATGCGCGTACAAAGAGATGGAGCCC
CGGTGGCCTCTTAAATTTCCCTTCTGCCACGGAGTTCGAAACCATCTACTCCACATGCGAG
GAGGCGGGTGGCAGCTGCAGCCCGAGTCCCCGTTCACTCAGGAAACGGAGACCTGTGAC
CACAGCAGGCTGACAGATGGACAGAATCTCCCGTAGAAAGGTTTGGTTTGAAGTCCCCGGG
GGCAGCAAACTGACATGTTGAATGATAGCATTTCACTCTGCGTTCTCCTAGATCTGAGCAA
GCTGTCACTTCTACCCCCACCGTGTATATACATGAGCTAACTTTTAAATTTGTACAAAA
GCGCATCTCAGATTCCAGACCTCGCCGATGACTTTTCTGAAGGCTGTGCTTTTCCCTCGC
CTTCTCTGAAGGTGCGATTAGAGCGAGTCCATGGAGCATCTTAACCTTTGCACTTTAGTTTT
TACAGTGAAGCTGAAGCTTTAAGTCTCATCCAGCATTTCTAATGCCAGGTTGCTGAGGGTAA
TTTTGAAGTAGATATATTACCTGGTCTGCTATCCTTAGTCAAACTCTGCGGTACAGGTAA
TTGAGAAATGTATACGGTACTTCCCTCCACACCATACGATAAAGCAAGACATTTATAACG
ATACAGAGTCACTATGTGGTCTCCCTGAAATAACGCATTGCAATTCAGTTCAGTGCAGTA
TATTTTTCTAAGTTTTGGAAAGCAGGTTTTTCTCTTAAAAAAATATAGACACGGTTCAT
AAATGATTTAGTCAGAAATCTAGACTGAAAGAACCTAAACAAAAAAATATTTTAAAGATA
TAAATATATGCTGTATGTTATGTAATTTATTTAGGCTATAATACATTTTCTATTTGCG
ATTTTCAATAAAATGCTCTAATAACAAAAA

FIGURE 4

MVKFPALTHYWPLIRFLVPLGITNIAIDFGEQALNRGIAAVKEDAVEMLASYGLAYSLMKFF
TGPMSEDFKNVGLVVFVNSKRDRTKAVLCMVVAGAIAAVFHTLIAYSDLGYIIINKLHHVDESV
GSKTRRAFLYLAAPPFMDAMAWTHAGILLKHKYSFLVGCASISDVIAQVVFAILLHSHLEC
REPLLIPILSLYMGALVRCCTTLCCLGYKNIHDIIPDRSGPELGGDATIRKMLSFWWPLALIL
ATQRISRPVNLVFSRDLGGSSAATEAVAILTATYPVGHMPYGWLTEIRAVYPAFDKNNPSN
KLVSTSNVTAAHIKKFTTFVCMALSLTLCFVMFWTFNVSEKILIDIIGVDFAPFAELCVVPLR
IFSFFPVPVTVRAHLTGWLMTLKKTFVLAPSSVLRIIVLIASLVVLPYLVGHGATLVGVGSL
AGFVGESTMVAIAACYVYRKQKKMENESATEGEDSAMTDMPTTEVTDIVEMRENE

Transmembrane domains:

amino acids 86-106, 163-179, 191-205, 237-253, 327-343, 357-374,
408-423, 431-445

FIGURE 5

CCTGACAGAAGTGCCCCGGAGCTGGGGGAGATNCAACATTAAGAAGATGCTGAGCTTCTGGT
GCCNTTTGGCTCTAATTCTGGCCACACAGAGAANCAGTCGGCCTATTGTCAACCTCTTTGTT
TCCCGGGACCTTGGTGGCAGTTCTGCAGCCACAGAGGCAGTGGCGATTTTGACAGCCACATA
CCCTGTGGGTCACATGCCATACGGCTGGTTGACGGAAATCCGTGCTGTGTATCCTGCTTTCG
ACAAGAATAACCCAGCAACAACTGGTGAGCACGAGCAACACAGTCACGGCGGCCACATC
AAGAAGTTCACCTTCGTCTGCATGGCTCTGTCACCTCACGCTCTGTTTCGTGATGTTTTGGAC
ACCCAACGTGTCTGNAAAAATCTTGATAGACATCATCGGAGTGGACTTTGCCTTTGCAGAAC
TCTGTGTGTTCTTTGCGGATCTTCTCTTCTTCCAGTTCAGTCACAGTGAGGGCGCAT
CTCACCGGGTGGCTGATGACACTGAAGAAAACCTTCGTC

TOTAL OF TOTALS

TGACGGGAATCCCGGGCTGGGTATCCTGGTTTNGACAAGATAAACCCCCAGCAANAATTTGGG
 GAGCAGGGGCAAAACAGTNACGGGCAGCCCATCAAGAAGTTTACCTTNGTTTGNATGNTC
 TGTCAACTCACGCTNTGTTTCGTGATGTTTTGGACACCCAAAGTGTTTGAGAAAAATTTTGAT
 AGACATNATCGGAGTGGANTTTGCCCTTGTGAGAANTTTGNGNTGTTCTCTTTGCGGATTTTCT
 CTTTTTCCCAAGTTCAGTCACAGNAGGGCGCATCTCACGGGNGGNTGATGACANTGAAG
 AAAACCTTTGTCTGTCGCCCCAGCATNTTTGGTCGCGATCATTTGTCCTNATNGCCAGCCTTGT
 GGTCTTACCTTACCTGGGGGTGCACGGTGCAGCCCTGGGCGTGGGTTCCCTCTGGCGGGCA

FIGURE 7

TATTCCCAGTTCCGGTCACGGGGAGGGCGCATNTCACCGGGTGGCTGANGACACTGAAGAAA
ACCTTNGTCCTTGCCCCCAGNTTTGTGNTGCGGATNATCGTCCTCATCGCCAGCCTNGTGGT
CCTACCCTACCTGGGGGTGCACGGTGAGAC

FIGURE 8

GGCCCGCGCCCGGCGCGGGCGCCGAAGCCGGGAGCCACCGCCATGGGGGCCTGCCTGGGA
GCCTGCTCCCTGCTCAGCTGCGCGTCTGCCTCTGCGGCTCTGCCCCCTGCATCCTGTGCAG
TGCTGCCCCGCCAGCCGCAACTCCACCGTGAGCGGCCTCATCTTCACGTTCTTCCTCTTCC
TGGGGGTGCTGTTGCCATATTATGCTGAGCCCCGGGCGTGAGAGTCACTCTACAAGCTG
CCCTGGGTGTGTGAGGAGGGGGCCGGGATCCCCACCGTCTGCAGGGCCACATCGACTGTGG
CTCCCTGCTTGGCTACCGCGCTGTCTACCGCATGTGCTTCGCCACGGCGGCCTTCTTCTCT
TCTTTTTACCCCTGCTCATGCTCTGCGTGAGCAGCAGCCGGGACCCCGGGCTGCCATCCAG
AATGGGTTTTGGTTCTTTAAGTTCCTGATCCTGGTGGGCCTCACCGTGGGTGCCTTCTACAT
CCCTGACGGCTCCTTACCAACATCTGGTTCTACTTCGGCGTCTGGGCTCCTTCTCTTCA
TCCTCATCCAGTGGTGCTGCTCATCGACTTTGCGCACTCCTGGAACAGCGGTGGCTGGGC
AAGGCCGAGGAGTGCATTTCCGTGCCTGGTACGAGGCCTCTTCTTCTTCACTCTCCTCTT
CTACTTGCTGTGATCGCGCCGTGGCGCTGATGTTTATGTACTACACTGAGCCAGCGGCT
GCCACGAGGGCAAGGTCTTTCATCAGCCTCAACCTCACCTTCTGTGTCTGCGTGTCCATCGCT
GCTGTCTGCCCAAGGTCCAGGACGCCAGCCCAACTCGGGTCTGCTGCAGGCCTCGGTCTAT
CACCTCTACACCATGTTTGTCACTGGTCAGCCCTATCCAGTATCCTGAAACAGAAATGCA
ACCCCATTTTGCCAACCCAGCTGGGCAACGAGACAGTTGTGGCAGGCCCCGAGGGCTATGAG
ACCCAGTGGTGGGATGCCCCGAGCATTTGTGGGCCTCATCATCTTCTCTCTGTGCACCCTCTT
CATCAGTCTGCGCTCCTCAGACCACCGGCAGGTGAACAGCCTGATGCAGACCGAGGAGTGCC
CACCTATGCTAGACGCCACACAGCAGCAGCAGCAGCAGGTGGCAGCCTGTGAGGCCCGGGCC
TTTGACAACGAGCAGGACGGCGTCACCTACAGCTACTCCTTCTTCCACTTCTGCCTGGTGCT
GGCCTCACTGCACGTCTATGATGACGCTACCAACTGGTACAAGCCCGGTGAGACCCGGAAGA
TGATCAGCACGTGGACCGCGTGTGGGTGAAGATCTGTGCCAGCTGGGCAGGGCTGCTCCTC
TACCTGTGGACCCTGGTAGCCCCACTCCTCCTGCGCAACCGCGACTTCAGCTTGAGGCAGCCT
CACAGCCTGCCATCTGGTGCCTCCTGCCACCTGGTGCCTCTCGGCTCGGTGACAGCCAACCT
GCCCCCTCCCCACCAATCAGCCAGGCTGAGCCCCACCCCTGCCCGAGCTCCAGGACCTG
CCCTGAGCCGGGCCTTCTAGTCGTAGTGCCCTTCAGGGTCCGAGGAGCATCAGGCTCCTGCA
GAGCCCCATCCCCCGCCACACCACACGGTGGAGCTGCCTCTTCTTCCCTCCTCCTCTGT
TGCCCACTCAGCATCTCGGATGAAAGGGCTCCCTTGCTCTCAGGCTCCACGGGAGCGGGG
CTGCTGGAGAGAGCGGGGAACCTCCACCACAGTGGGGCATCCGGCACTGAAGCCCTGGTGTT
CCTGGTCAGCTCCCCAGGGGACCTGCCCCCTTCTGGACTTCGTGCCTTACTGAGTCTCT
AAGACTTTTTCTAATAAACAGCCAGTGCCTGTAAAAAAA

FIGURE 9

MGACLGACSLSCASCSCGSA PCILCSCCPASRNSTVSRLIFTFFLFLGVLVSIIMLSPGVE
SQLYKLPWVCEGAGIPTVLQGHIDCGSLLGYRAVYRMCFATAAFFFFFFFFLLMLCVSSSRD
PRAAIQNGFWFFKFLILVGLTVGA FYIPDGSFTNIWFYFGVVG SFLFILIQLVLLIDFAHSW
NQRWLGKAEECD SRAWYAGLFFFTLLFYLLSIAAVALMFMYYTEPSGCHEGKVFI SLNLTFC
VCVSI AAVLPKVQDAQPN SGLLQASVITLYTMFVTWSALSS IPEQKCNPHLPTQLGNETVVA
GPEGYETQWW DAPSIVGLIIFLLCTLFISLRSSDHRQVNSLMQTEECPPMLDATQQQQQQVA
ACEGRAFDNEQDGV TYSYSFFHFCLVLASLHVMMTLTNWYKPGETRMISTWTAVVWKICAS
WAGLLLYLWTLVAPLLLRNRDFS

Signal sequence:

amino acids 1-20

Transmembrane domains:

amino acids 40-58, 101-116, 134-150, 162-178, 206-223, 240-257,
272-283, 324-340, 391-406, 428-444

FIGURE 10

GAGCGAGGCCGGGACTGAAGGTGTGGGTGTCGAGCCCTCTGGCAGAGGGTTAACCTGGGTC
AAATGCACGGATTCTCACCTCGTACAGTTACGCTCTCCGCGGCACGTCCGCGAGGACTTGA
AGTCCTGAGCGCTCAAGTTTGTCCGTAGGTCGAGAGAAGGCCATGGAGGTGCCGCCACCGGC
ACCGCGGAGCTTTCTCTGTAGAGCATTGTGCCTATTTCCCCGAGTCTTTGCTGCCGAAGCTG
TGACTGCCGATTTCGGAAGTCCTTGAGGAGCGTCAGAAGCGGCTTCCCTACGTCCCAGAGCCC
TATTACCCGAATCTGGATGGGACCGCCTCCGGGAGCTGTTTGGCAAAGATGAACAGCAGAG
AATTTCAAAGGACCTTGCTAATATCTGTAAGACGGCAGCTACAGCAGGCATCATTTGGCTGGG
TGTATGGGGGAATACCAGCTTTTATTCATGCTAAACAACAATACATTGAGCAGAGCCAGGCA
GAAATTTATCATAACCGGTTTGATGCTGTGCAATCTGCACATCGTGTGCCACACGAGGCTT
CATTCGTTATGGCTGGCGCTGGGGTTGGAGAACTGCAGTGTTTGTGACTATATTCAACACAG
TGAACACTAGTCTGAATGTATACCGAAATAAGATGCCTTAAGCCATTTTGTAAATTGCAGGA
GCTGTACCGGAAGTCTTTTATGATAAACGTAGGCCTGCGTGGCCTGGTGGCTGGTGGCAT
AATTGGAGCCTTGCTGGGCACCTCCTGTAGGAGGCTGCTGATGGCATTTCAGAAGTACGCTG
GTGAGACTGTTTCAGGAAAGAAAACAGAAGGATCGAAAGGCACCTCCATGAGCTAAAACTGGAA
GAGTGGAAGGCAGACTACAAGTTACTGAGCACCTCCCTGAGAAAATTGAAAGTAGTTTACG
GGAAGATGAACCTGAGAATGATGCTAAGAAAATTGAAGCACTGCTAAACCTTCTAGAAACC
CTTCAGTAATAGATAAAACAAGACAAGGACTTGAAAGTGTCTGAACTTGAAACTCACTGGAGA
GCTGAAGGGAGCTGCCATGTCCGATGAATGCCAACAGACAGGCCACTCTTTGGTCAGCCCTGC
TGACAAATTTAAGTGTCTGGTACCTGTGGTGGCAGTGGCTTGCTCTTGTCTTTTCTTTTCTT
TTTAACTAAGAATGGGGCTGTTGTACTCTCACTTTACTTATCCTTAAATTTAAATACATACT
TATGTTTGTATTAATCTATCAATATATGCATACATGGATATATCCACCACCTAGATTTTAA
GCAGTAAATAAAACATTTTCGCAAAAGATTAAAGTTGAATTTTACAGTTT

FIGURE 11

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA23318

><subunit 1 of 1, 285 aa, 1 stop

><MW: 32190, pI: 9.03, NX(S/T): 2

MEVPPPAPRSFLCRALCLFPRVFAAEAVTADSEVLEERQKRLPYVPEFYYPESGWDRLRELF
GKDEQQRISKDLANICKTAATAGIIGWVYGIPAFIHAKQQYIEQSQAETIYHNRFDAVQSAH
RAATRGFIIRYGWRWGWRRTAVFVTIFNTVNTSLNVYRNKDALSHFVIAGAVTGSGLFRINVGLR
GLVAGGIIGALLGTPVGGLLMAFQKYAGETVQERKQKDRKALHELKLEEWKGRQLQVTEHLPE
KIESSLREDEPENDAKKIEALLNLPRNPSVIDKQDKD

Important Features:

Signal Peptide:

amino acids 1-24

Transmembrane domains:

amino acids 76-96 and 171-195

N-glycosylation site:

amino acids 153-156

THEORY OF THE EARTH

CGGAAGTCCCTTGAGGAGCGTFCAGAAGCGGCTTCCCTACGTCCCAGAGCCCTATTACCGGA
ATCTGGATGGGACCGCTCCGGGAGCTGTTTGGCAAGATGAACAGCAGAGAATTTCAAAGGA
CCTTGCTAATATCTGTAAGACGGCAGCTACAGCAGGCATCATTTGGCTGGTGTATGGGGGAA
TACCGACTTTTATTATCATGCTAAACAACAATCATTGAGCAGAGCCAGGCAGAAATTTATCAT
AACCGGTTTGTATGCTGTGCAATCTGCACATCGTGTCTGCCACAGAGGCTTCATTCTGTTATG
GCTGGCGCCGAACC

FIGURE 13

TCAAGTTTGTCGGTAGGTCGAGAGAAGGCCATGGAGGTGCCGCCACCGGCACCGCGGAGCTT
TTTTCTGTAGAGCATTGTGCCTATTTCCCCGAGTTTTTGCTGCCGAAGCTGTGACTGCCGAT
TCGGAAGTCCTTGAGGAGCGTCAGAAGCGGCTTCCCTACGTCCCAGAGCCCTATTACCCGGA
ATTTGGATGGGACCGCCTCCGGGAGCTGTTTGGCAAAGATGAACAGCAGAGAATTTCAAAGG
ACCTTGCTGATATNTGTAAGACGGCAGCTACAGCAGGCATCATTTGGCTGGGTGTATGGGGGA
ATACCAGCTTTTATTCA TGNTAAACAACAATACATTGAGCAGAGCCAGGCAGAAATTTATNA
TAACC

1017191.102401

NOTES

GAGCGCCGCGCGCGCGCGCGCGCGCACTGCAGCCCCAGGCCCCGGCCCCCAACCAAGCTCT
GCGTTGCTGCCCCGCGCTGGGCCAGGCCCCAAAGGCAAGGACAAGCAGCTGTCAAGGAACCT
CCGCCGAGTTCGAATTTACGTGCAGCTGCCGCGCAACCAAGGTTCCAAGATGGTTTTCGCGGG
GCTTCGCGTGTTCGAAGAACTGCCTGTGCGCCCCCAACCTGCCTTTACACCTTGGTTAGTCTG
CTGCTAAATTGGAATTTGCTGCGTGGGGCATTGGCTTCGGGCTGATTTCCAGTCTCCGAGTGGT
CGGCGTGTTCATTGCAGTGGGCATCTTTCTGTTCTGATTGCTTTAGTGGGTCTGATTGGAG
CTGTAAAAACATCATCAGGTGTTGCTATTTTTTTTATATGATTATCTGTTACTTGTATTATT
GTTACAGTTTTCTGTATCTTGCGCTTGTTTAGCCCTGAACCAGGAGCAACAGGGTCAGCTTCT
GGAGGTTGGTTGGAAACAATACGGCAAGTGCTCGAAATGACATCCAGAGAAATCTAAACTGCT
GTGGGTTCCGAAGTGTTAAACCCAAATGACACCTGTCTGGCTAGCTGTGTTAAAGTGACCAC
TCGTGCTCGCCATGTGCTCCAATCATAGGAGAATATGCTGGAGAGGTTTTGAGATTTGTTGG
TGGCATTGGCCTGTTCTTCAGTTTTACAGAGATCCTGGGTGTTTGGCTGACCTACAGATAACA
GGAACCGAAGAGACCTCCGCGCGAATCCAGCTAGTGATCTCTTTGATGAGAAAACAAGGAAGAT
TCTCTTTCGTATTATGATCTTGTTCACCTTTCTGTAATTTCTGTTAAGCTCCATTGTTGCCAGT
TTAAGGAAGGAACACIATCTCGAAAAGTACCTTATGTATAGTGGAAATATATATTTTTACT
CTATGTTTCTCTACATGTTTTTTCTTTCCGTTGCTGAAAAATATTTGAAACTTGTGGTCTC
TGAAGCTCGGTGGCACCTGGAATTTACTGTATTCAATGTGCGGCACGTGCCACTGTGGCCTT
TCTTAGCATTTTTACCTGCAGAAAAACTTTGTATGGTACCACTGTGTTGGTTATATGGTGAA
TCTGAACGTACATCTCACTGGTATAATTATATGTAGCACTGTGCTGTGTAGATAGTTCCTAC
TGAAAAAAGAGTGGAATTTATTAATAACAGAAAGTATGAGATCCTGTTATGTTAAGGGAAA
TCCAAATTCCTAATTTTTTTTTGGTCTTTTAGGAAAGATTGTTGTGGTAAAAAGTGTTAGTA
TAAAAATGATAATTTACTTGTAGTCTTTATGATTACACCAATGTATCTAGAAATAGTTAT
GTCTTAGGAAATGTGGTTTAATTTTTGACTTTTACAGGTAAGTGCAAGAGGAGAAGTGTTTT
CATGAAATGTTCTAATGTATAATAACATTTACCTTCAGCCTCCATCAGAATGGAACGAGTTT
TGAGTAATCAGGAAGTATATCTATATGATCTTGATATTGTTTTATAATAATTTGAAGTCTAA
AAGACTGCATTTTTAAACAAGTTAGTATTAAATGCGTTGGGCCACGTAGCAAAAAGATATTTG
ATTATCTTAAAAATTGTTAAATACCGTTTTTCATGAAATTTCTCAGTATTGTAAACAGCAACTT
GCTCAAACTAAGCATATTTGAAATATGATCTCCATAAATTTGAAATGAAATCGTATTTGTGTG
GTCTGTATATTTCTGTTAAAAAATTAAGGACAGAAACCTTTCTTTGTGTATGCATGTTTGA
ATTAAGAAGAAAGTAATGGAAG

FIGURE 15

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA39979

><subunit 1 of 1, 204 aa, 1 stop

><MW: 22147, pI: 8.37, NX(S/T): 3

MVCGGFACSKNCLCALNLLYTLVSLLLIGIAAWGIGFGLISSLRVVGVIIVGIFLFLIALV
GLIGAVKHHQVLLFFYMIILLLVFIVQFSVSCACALNQEQQGQLLEVGVWNNTASARNDIQR
NLNCCGFRSVNPNDTCLASCVKSDHSCSPCAPIIGEYAGEVLRFGVGIGLFFSFTBILGVWL
TYRYRNQKDPRANPSAFL

Signal Peptide:

amino acids 1-34

Transmembrane domains:

amino acids 47-63, 72-95 and 162-182

FIGURE 16

TGATTGGAGCTGTAAAAAANTCTTCAGGTGTTGTNATTTTTTTATATGATTATTCTGTAANT
TGTATTTATTGTTTCAGTTTTNTGTATCTTGCGCTTGTTAGCCNTGAACCAGGAGCAACAGG
GTCAGNTTNTGGAGGTTGGTTGGAACAATACGGCAAGTGCTCGAAATGACATCCAGAGAAAT
NTAAACTGCTGTGGGTTCCGAAGTGTTAACCCAAATGACACCTGTNTGGCTAGCTGTGTAA
AAGTGACCACTNGTGCTCGCCATGTGCTCCAATCATAGGAGAATATGCTGGAGAGGTTTTGA
GATTTGTTGGTGGCATTGGCCTGTTNTTCAGTTTTTACAGAGATCCTGGGTGTTTGGCTGACC
TACAGATACAGGAACCAG

FIGURE 17

AATCCCAAATCCCCAATTTTTTGGNCTTTTTAGGGAAAGATGTGTTGTGGTAAAAAGTGT
TAGTATAAAAAATGATAATTTACTTGTAGTCCTTTATGATTACACCAATGTATTCTAGAATAG
TTATGTCTTAGGAAATTGTGGTTAATTTTTGACTTTTACAGGTAAGTGCAAAGGAGAAGTG
GTTTCATGAAATGTTCTAATGTATAATAACATTTACCTTCAGCCTCCCATCAGAATGGAACG
AGTTTTGAGTAATCCAGGAAGTATATCTATATGATCTTGATATTGTTTTATATAATTTGAAG
TCTAAAAGACTGCATTTTTAAACAAGTTAGTATTAATGCGTTGGCCACGTAGCAAAAAGAT
ATTTGATTATCTTAAAAATTGTTAAATACCGTTTTTCATGAAAGTTCTCAGTATTGTAAACAGC
AACTTGTCAAACCTAAGCATATTTGAATATGATCTCCATAATTTGAAATTGAAATCGTATT
GTGTGGAGGAAATGGCAATCTTATGTGTGCTGAAGGACACAGTAAGAGCACCAAGTTGTGCC
CCACTTGC

FIGURE 18

ATGATTATTCTGTTACTTGTATTTATTGTTTCAGTTTATGGTATCTTGCGCTTGTTAGCCC
CTGAAACCAGGAGCAACAGGGNNCAGCTTCCTGGAGGTTGGTTGGCAACAATCACGGCCAAG
TGACTCCGCAAATGACATCCCAGAGAAATCCTAAACTGCTGTGGGTTCCGAAGTGTTAACCC
AAATGACACCTGTCTGGCTNGCTGTGTAAAGTGACCACTCGTGCTCGCCATGTGCTCAA
TCATAGGAGAATATGC

10017191.102402

FIGURE 19

CAGTCACCATGAAGCTGGGCTGTGTCTCATGGCCTGGGCCCCTACCTTTCCCTTTGGTGTG
CTCTGGGTGGCCAGATGCTACTGGCTGCCAGTTTTGAGACGCTGCAGTGTGAGGGACCTGT
CTGCACCTGAGGAGAGCAGCTGCCACACGGAGGATGACTTGACATGATGCAAGGGAGCTGGCT
TCCAGGTCAAGGCCCTACACTTTTCAGTGAACCTTCCACCTGATTGTGTCTATGACTGGCTG
ATCTCCAAGGTCAGCCAAGCCAGTTTTTTGAAGGGGACCTGCTGGTCTGCGCTGCCAGGC
CTGGCAAGACTGGCCACTGACTCAGGTGACCTTCTACCGAGATGGCTCAGCTCTGGGTCCCC
CCGGCCCTAAACAGGGAATTCTCCATCACCGTGGTACAAAAGGCAGACAGCGGGCACTACCAC
TGCAGTGGCATCTTCCAGAGCCCTGCTCCTGGGATCCAGAAACAGCATCTGTTGTGGCTAT
CACAGTCCAAGAACTGTTTTCCAGCGCCAATTCTCAGAGCTGTACCCCTCAGCTGAACCCCAAG
CAGGAAGCCCCATGACCCCTGAGTTGTGAGACAAAGTTGCCCTGACAGAGGTCAGCTGCCCGC
CTCCTCTTCTCCTTCTACAAGGATGGAAGGATAGTGCAAAGCAGGGGGCTCTCCTCAGAATT
CCAGATCCCCACAGCTTCAGAAGATCACTCCGGGTCACTACTGGTGTGAGGCAGCCACTGAGG
ACAACCAAGTTTGAAACAGAGCCCCCAGCTAGAGATCAGAGTGCAGGGTGCTTCCAGCTCT
GCTGCACCTCCACATTTGAATCCAGCTCCTCAGAAATCAGCTGCTCCAGGAACCTGCTCCTGA
GGAGGCCCTGGGCTCTGCTCCGCGCCAACCCCATCTTCTGAGGATCCAGGCTTTTCTT
CTCCTCTGGGGATGCCAGATCCTCATCTGTATCACCAGATGGGCCTTCTTCTCAAACACATG
CAGGATGTGAGAGTCTCCTCGGTCACTGCTCATGGAGTTGAGGGAATTATCTGGCCACCA
GAAGCCTGGGACCACAAAGGCTACTGCTGAATAGAAGTAAACAGTTCCATCCATGATCTCACT
TAACCACCCCAATAAATCTGATTCTTTATTTTCTCTTCTGCTGTCACATATGCATAAGTA
CTTTTACAAGTTGTCCAGTGTTTTGTTAGAATAATGTAGTTAGGTGAGTGTAATAAATTT
ATATAAAGTGAGAATTAGAGTTTAGCTATAATTGTGTATTCTCTCTTAAACACAACAGAATTC
TGCTGTCTAGATCAGGAATTTCTATCTGTTATATCGACCAGAATGTTGTGATTTAAAGAGAA
CTAATGGAAGTGGATTGAATACAGCAGTCTCAACTGGGGGCAATTTTGCCCCCAGAGGACA
TTGGGCAATGTTTGAGACATTTTGGTCATTATACTTGGGGGGTGGGGGATGTTGGGATGT
GTGTCTACTGGCATCCAGTAAATAGAAGCCAGGGGTGCCGCTAAACATCCTATAATGCACAG
GGCAGTACCCACAAACGAAAAATAATCTGGCCCAAAATGTGAGTTGTAAGTGTGAGTTGAGAAA
CCCCAGCCTAATGAAACCCTAGGTGTTGGGCTCTGGAATGGGACTTTGTCCCTTCTAATTAT
TATCTCTTCCAGCCTCATTCAGCTATTTCTACTGACATACCAGTCTTTAGCTGGTGCTATG
GTCTGTTCTTTAGTTCTAGTTTGTATCCCTCAAAGCCATTATGTTGAAATCCTAATCCCC
AAGGTGATGGCATTAAAGAGTGGGCTTTGGGAAGTGATTAGATCAGGAGTGCAGAGCCCTC
ATGATTAGGATTAGTGCCCTTATTTAAAAAGGCCCCAGAGAGCTAACTCACCCCTTCCACCAT
ATGAGGACGTGGCAAGAAGATGACATGTATGAGAACCAAAAAACAGCTGTGCGCAACACCG
ACTCTGTCTGGCTTGATCTTGAACCTCCAGCCTCCAGAACTATGAGAAATAAATTTCTGG
TTGTTTGTAGCCTAA

1001794.102401

```
><subunit 1 of 1, 359 aa, 1 stop
```

MKLGCVLMAWALYLSLGVLVVAQMLLAASFETLQCEGPVCTEESSCHTEDDLTDAREAGFQV
 KAYTFSEPFHLIVSYDNLILQGPAKPVFEGDLVLVRCQAWQDWPLTQVTFYRDGSALGPPGP
 NREFSITVQKADSGHYHCSGIFQSPGPGIPETASVVAITVQELFPAPILRAVPSAEPQAGS
 PMTLLSCQLEKPLRQSAARLLFSFYKDNPRIVQSRGLSSEFQIPTASEDSHGSYWCEAATEDNQ
 VWKQSPQLEIRVQGASSAAPPTLNPAQKSAAPGTAEAPAGPLPPPTPSSDPGFSFSL
 GMPPDPLYHQMGLLLHKHMDPVRVLLGHLLMRELSGHOKPGTTKATAE

amino acids 1-17

amino acids 12-33

amino acids 353-355

FIGURE 21

CCCACGCGTCCGCCACGCGTCCGCCACGCGTCCGCCACGCGTCCGGGCCACCAGAAGTT
TGAGCCTCTTTGGTAGCAGGAGGCTGGAAAGAAAGGACAGAAGTAGCTCTGGCTGTGAATGGGG
ATCTTACTGGGCCTGCTACTCTCTGGGGCACCTAACAGTGGACACTTATGGCCGTCCCATCTCT
GGAAGTGCCAGAGAGTGTAAACAGGACCTTGGAAGGGGATGTGAATCTTCCCTGCACCTATG
ACCCCTTGCAAGGCTACACCCAAGTCTTGGTGAAGTGGCTGGTACAACGTGGCTCAGACCCT
GTCACCATCTTTCTACGTGACTCTTCTGGAGACCATATCCAGCAGGCAAAGTACCAGGGCCG
CCTGCAATGTGAGCCACAAGGTTCCAGGAGATGTATCCCTCCAATTGAGCACCTTGGAGATGG
ATGACCCGAGGCCACTACAGTGTGAAGTCACTTGGCAGACTCTGTATGGCAACCAAGTCGTG
AGAGATAAGATTACTGAGTCCGTGTCCAGAACTCTCTGTCTCCAAGCCACAGTGACAAC
TGGCAGCGGTTATGGCTTCAAGGTGCCCCAGGGAATGAGGATTAGCCTTCAATGCCAGGCTC
GGGGTTCTCTCCCATCAGTTATATTGGTATAAGCAACAGACTAATAACAGGAACCCATC
AAAGTAGCAACCTAAGTACCTTACTCTTCAAGCCTGCGGTGATAGCCGACTCAGGCTCCTA
TTTCTGCACTGCCAAGGGCCAGGTTGGCTCTGAGCAGCAGCGACATTGTGAAGTTTGTGG
TCAAAGACTCTCAAAGCTACTCAAGACCAAGACTGAGGCACCTACAACCATGACATACCCC
TTGAAAGCAACATCTACAGTGAAGCAGTCTTGGGACTGGACCATGACATGGATGGCTACCT
TGAGAGAGACCAGTGTGGGCCAGGAAAGAGCCTGCTGTCTTTGCCATCATCCTCATCATCT
CCTTGTGCTGTATGGTGGTTTTTACCATGGCCTATATCATGCTCTGTGCGAAGACATCCCAA
CAAGAGCATGTCTACGAAGCAGCCAGGTAAAGAAAGTCTCTCCTCTTCCATTTTTGACCCCGT
CCCTGCCCTCAATTTTGATTACTTGGCAGGAAATGTGGAGGAAGGGGGGTGTGGCACAGACCC
AATCCTAAGGCCGAGGCGCTTCAAGGTCAAGACATAGCTGCCTTCCCTCTCAGGCACCTT
CTGAGGTTGTTTTGGCCCTCTGAACACAAAGGATAATTTAGATCCATCTGCCTTCTGCTTCC
AGAATCCCTGGGTGTAGGATCTTGATAATTAATTGGCAAGAATTGAGGCAGAAGGGTGGGA
AAGCAGGACCACAGCCCAAGTCCCTTCTTATGGGTGGTGGGCTCTTGGGCCATAGGGCAC
TGCCAGAGAGGCCAACGACTCTGGAGAAACCATGAGGGTGGCCATCTTCGCAAGTGGCTGCT
CCAGTGATGAGCCAACTTCCAGAATCTGGGCAACAACACTACTCTGATGAGCCCTGCATAGGA
CAGGAGTACCAGATCATCGCCAGATCAATGGCAACTACGCCCGCCTGCTGGACACAGTTCC
TCTGGATTATGAGTTTCTGGCCACTGAGGGCAAAAGTGTCTGTTAAAAATGCCCATATTAGGC
CAGGATCTGCTGACATAATTGCCTAGTCAGTCCCTTGCCTTCTGCATGGCCTTCTTCCCTGCT
ACCTCTCTTCTCTGGATAGCCAAAGTGTCCGCTACCAACACTGGAGCCGCTGGGAGTCACT
GGCTTTGCCCTGGAAATTTGCCAGATGCATCTCAAGTAAGCCAGCTGCTGGATTTGGCTCTGG
GCCCTTCTAGTATCTCTGCCGGGGGCTTCTGGTACTCCTCTCTAAATACAGAGGGAAGATG
CCCATAGCACTAGGACTTGGTCATATGCCTACAGACACTATTCAACTTTGGCATCTTGCCA
CCAGAAGACCCGAGGGAGGCTCAGCTTCCGACTCAGAGGACCAGCTATATCCAGGATCAT
TTCTCTTTCTCAGGGCCAGACAGCTTTTAATTGAAATTGTTATTACAGGCCAGGGTTCA
GTCTTGCTCTCTCCACTATAAGTCTAATGTTCTGACTCTCTCCTGGTGCTCAATAAATATCTA
ATCATACACG

FIGURE 22

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45416

><subunit 1 of 1, 321 aa, 1 stop

><MW: 35544, pI: 8.51, NX(S/T): 0

MGILLGLLLLLGHLTVDTYGRPILEVPESVTGPWKGDVNLPCITYDPLQGYTQVLVKWLVQRGS
DPVTIFLRDSSGDHIQQAKYQGRLHVSHKVPGDVSLQLSTLEMDDRSHYTC EVTWQTPDGNQ
VVRDKITELRVQKLSVKPTVTGTSGYGF TVPQGM RISLQCQARGSPPI SYIWYKQQTNNQE
PIKVATLSTLLFKPAVIADSGSYFCTAKGQVGSEQHSDIVKFVVKDSSKLLKTKTEAPTMT
YPLKATSTVKQSWDWT TDM DGYLGETSAGPGKSLPVFAI ILIISLCCMVVFTMAYIMLCRKT
SQQEHVYEAAAR

Signal Sequence:

amino acids 1-19

Glycosaminoglycan attachment site:

amino acids 149-152

Transmembrane domain:

amino acids 282-300

FIGURE 24

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45419

><subunit 1 of 1, 373 aa, 1 stop

><MW: 41281, pI: 8.33, NX(S/T): 3

MSLLLLLLVSYVGTGLGTHTEIKRVAEEKVTLPCHHQLGLPEKDTLDIEWLLTDNEGKQKV
VITYSSRHVYNNLTEEQKGRVAFASNFLAGDASLQIEPLKPSDEGRYTCVKVNSGRYVWSHV
ILKVLVRPSKPKCELEGELTEGSDLTLCCESSSGTEPIVYVWQRIREKEGEDERLPPKSRID
YNHPGRVLLQNLTMSYSGLYQCTAGNEAGKESCVVRVTVQYVQSIGMVAGAVTGIVAGALLI
FLLVWLLIRRKDKERYEEEEERPNEIREDAEAPKARLVKPSSSSSGSRSSSGSSSTRSTANS
ASRSQRTLSTDAAPQFGLATQAYSLVGPEVRGSEPKKVHHANLTKAETTPSMIPSQSRAFQTV

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 232-251

GTGGCTCTCTTTGCTCTCTCTCTCGGACGATCTCTCCCTGCTTCTCTCTCTCGACGCGCTGTGCGAGAGAGACACCCGGA
AGACGGGGCTGACGATCGCGGGGGCTTCTCCCGCTGGGCGGCCTCGCGCTGTGGCAGGTGCTGAGGCGCTCTAG
AGCTCTGTTGCGCGCTCTCTCTCTGCGCGGGCGAGCAGCAGTACACATGGGGTGTGTGGAGGTAGATGGGCTCCCG
GCCCGGGGAGCGGGCGGTGGATGCGCGGCTGGGACAGAGACGCGCGGATTTTCAGCTGTGCGCGCGCGCGCGGGGG
CCCCGCGAGTCTCCCGGTTCAGCCATTTGGGACCTTCTCCGAGCAGGACCGGCTCGCTCTCTGCGAGCGATCG
GCCCGCGGAGCGACAGCAGATGATCGGGGGTCCCTTTCTCTGTGGATTCTCTTAGCACAACAGCAGCTCAG
CGAAGACGAGAGGGCTCGAATTCATTTGGCACATACCGCATGTGACCGTGCCACCGCGAGGTGTCTAACTGT
GACAAGTGTCCAGACGAGAACTGATGTCTTGAGACTTTGACCAACAACGCTCGCGCTGTGACGAGATGCCCT
GTGGGACGTTTACCGAGCATGAGAATGGCATGAGAAATGCATGATGTAGTCAGCACTGCCCATGGCCAAATG
ATTGAGAATACTTGTGTGCTGCTGATCTGACCGAGAATGCATTTGCCCACTGGGATGTTCAGTCTTAAGCT
ACCTGTGCGCCCCATCGGTTGTGTCTGTGGTGTGGGTGTGGGAAGAAAGGACAGAGACTGAGGATGTGCGCT
TGTAAGCACTGTGCTGGGGTACCTTTCTCAGATGTGCTTTCTAGTGTGATAGAAATGCAAGACATACACAGACTGT
CTAGATCGAACTGCTGGTGGTGTACAGCGCGGGACAGGAGACAGCAACACTGTGTGGACACTCCGCTCTCTC
TCGAGCTCCACCTCACTTCTCGCTGGCAGCAAGTTCATTCACGCTTCGAGCAGATGAAACCCATGAGTCCCT
TCTCCACTTATGTTCCAAAGGCTGAACCTAACAGAACTCAACTTCTGCTCTGTGTAGACAGAGGATCTG
AGTAGAATCCAGGAAAGGACAGTCCCTGACAAACAACGCTCAGCAGAGGGGAGGAAGAGCTGAACAAGACCTCT
CCAAACCTTCAGGTAGTCAACACAGCAGGCGCCACACAGCAACATCTGAAGTGTCTGCGGTCTCATGAG
GCCACTGGGGGCGAAGATCCGACACGCCCTCAAGSGCCCAAGGCGGAGCTCTTAGACGAACCTACACAAG
CACTTTTGACATCAATGAGACTTCCTTCTGGATGATGTGTGCTTTCTGCTGTGTGCTGTGGTGTATGTGTGT
TCGAGTATCCGGAAGAGCTCGAGGACTCGAAGAAAGGGGCGCGCGAGTCCAGTGCATTTGTGGAAAGAGCGA
GGGCTAGAAGAACTCATGACTCAACCAACGAGCCGAGCAATGATCTATCTAGTCAATGGCCATGTGTATGAT
ATCTGAGGCTTTAGCAGCCAACTGGGAGAGCTCGAGGATGATCAAGTATCTCAAGTTCTTTCAGTATGCCATBAG
AGGGAGTGTGCTGCTTTCTTCCATGGGTACACGCGCGACACGAGCGGCGCTACGAGCTCTGCACGATCGACCT
CTCGGGGCGCGGAGCGAGCTCGCGCGAGTCAATTAGCGGCTGTGCGCCAGCAGCGGAGAGAAAGATGTGTGGAG
AGATTTCTGGGCTGTATGGAAGACACACCGCACTGAAAGATGACAAACTAGCTCTCCGATAGCGCCGACCGCG
CTTAGCCGAGCCCCATCCCCAGCCCAACGCGAAATTTGAGAACTCCGCTCTCTAGCGGTGGAGCGTCCCCA
CAGGACAGAAACAGGGGCTCTCTGTGGTAGTGTCAAGAGCGGCTCTCGCTGTGATCTACATACCGCGGCTCT
TCGCGCTGAGCAGGAACGGTTCTGTTTATTAACGAAAGAGAGGACAGCAGTGTGCGCGAGGTACGCGCTGAC
CCCTGTGCTGCGGCTCTTTGATGACGATCTCCACTTTCTAAACTCTGAGGAGCTCGGGGTGATTGAAGAG
ACTTCCGAGCTGGAGGACAACTTAGACGGCTATCGAAATTTAGGAGTCAAGAGCCAGAGAGCCAGCCAGCC
CTCTCGAGCTGTTTATGAGCATCTCTTCGATCGCTGTGTAAGACATAGGATATCGCATTTCTGAAATTAATCA
ATTATGTGCGAGGTGTGTTTTAAATTTCTCTGTGTTTAAAGAAATGATGCGGAGTGTGTGTGTTGTTGTT
GTGTGTGTGTGTGTGTGTGTGTGTGTGTTTAAACAGAGAATATGGCCAGTGTGATGCTTTCTCTCTCTCT
TCTCTCTCTTTTTTTTTTAAATTAACCTCTCGGGAAGTTGGTTTAAAGAGCTTGGCAGGTGTAATCTGTGTGAA
ATACCCACACACATAAGATTTTTTAAGTCCCATATTTCTCTCAATTTGCTTTATGATTTTTCAGATTTATCTG
TGCACTTTAAATTTTACTTAACATACATAGATGAGTGTGACTTTTCCACACACGAGTTATGGAGGCTCTTAACT
TCTTAAAGATTAATATGGCATCTGTGAATCTTGAATGACGATTTATGTGCTCTTAACATACACACCTACTTTTT
AAAAACAATATATATCTATTTTTATTTATTTGTTGCTTTTAAATATTTCAAAGATTAAGAAATTTAAGA
CCCCATTGAGTTCTGATGATGAATCAACTTTGAGTTATCTTTAAATATCTGTGTATGATCTATATTCATGCT
CTGAAGTTGACCACTAATTCGATGTTGATGTTTACCTCGGACACCGGTGAGAAGCTGTGATTACTGATGAC
TCTTCTTATGCTAATATGCTCTGGGCTGGGAAGATGAATCCTCAGGCACTCAGGATTTGCTATTTAAGTGCGTT
GCAACTGGGGCAACAAAGAACTGAACTCACTTCACTTTAGGATTTAGGCTGTGTGAAACAACATGCTGCACTT
GGAAGGTCAAAATCAAGTCCGAGTGGCGGCTTTCCATAGAGAACTTTGGCCAGCTTTGCTTTTAAAGATGTCTGT
TTTTTATATACATAATGCAATAGAGTCCAACTGTGCTCTCAAGGCTTTGCTCTGTGGGATTCTGCTCAACAT
ACTTTTATTAATAATGAGTGTGCACTGTGAAGACCTTTGTCTGATATTTGCAACTGTGCCATTTACAAAT
TACCTTCTAACTGCTTGTGCGAGTCCCAATCAAGAGGTGGCTGAGTCCCTTTGTGTGGGTGGGTTTGTGG
GTAGTGTGGAGGACCGGATTCAGAAAAATGCCCTTCAGTGTATCAATTTATTAATAAACATTAGGCTTTGTGTA
AAAAAATAA

FIGURE 26

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52594

><subunit 1 of 1, 655 aa, 1 stop

><MW: 71845, pI: 8.22, NX(S/T): 8

MGTSPSSSTALASCSRIARRATATMIAGSLLLLGFLSTTTAQPEQKASNLIQTYRHVDRATG
QVLTCDKCPAGTYVSEHCTNTSLRVCSSCPVGTFTRHENGIEKCHDCSQPCPWPMIEKLPCA
ALTDRECTCPPGMFQSNATCAPHTVCPVGWGVRRKGTETEDVRCKQCARGTFSDVPSSVMKC
KAYTDCLSQNLVVIKPGTKETDNVCGTLPSPFSSSTSPSPGTAIFPRPEHMETHEVPSSTYVP
KGMNSTESNSSASVRPKVLSSIQEGTVPDNTSSARGKEDVNKTLPNLQVVNHQQGPHHRHIL
KLLPSMEATGGEKSSTPIKGPKRGHPRQNLHKHFDINEHLPWMIVLFLLLVLVIVVCSIRK
SSRTLKKGPRQDPSAIVEKAGLKKSMPTPTQNREKWIYYCNGHGIDILKLVAQVGSQWKDIY
QFLCNASEREVAFAFSNGYTADHERAYAALQHWITIRGPEASLAQLISALRQHRRNDVVEKIRG
LMEDTTQLETDKLALPMSPLSPSPPIPSNAKLENSALLTVEPSQDKNGFFVDESEPLL
RCDSTSSGSSALSRNGSFITKEKDTVLRQVRLDPCDLQPIFDDMLHFLNPEELRVIEEIPQ
AEDKLDRLFEIIGVKSQEASQTLSDSVYSHLPDLL

Signal sequence:

amino acids 1-41

Transmembrane domain:

amino acids 350-370

FIGURE 27

ATGGAAGCCAGTAACACTGTGGCCTACTATCTCTTCCGTGGTCCACTACATTITTTGGGA
CTCGGAATTATGAGGTAGAGGTGGAGGCGGAGCCGGATGTGAGAGGTCCTGAAATAGTCAC
CATGGGGGAAATGATCCGCTGCTGTTGAAGCCCCCTTCTCATTCCGATCGCTTTTGGCC
TTGATGATTTGAAAATAAGTCCCTGTTGCACAGATGCAGATGCTGTTGCTGCACAGATCTG
TCACTGCTGCCATTGAAGTTTTTCCAATCATCGTCATTGGGATCATTGCATTGATATTAGC
ACTGGCCATTGGTCTGGGCATCCACTTCGACTGCTCAGGGAAGTACAGATGTCGCTCATCCT
TTAAGTGTATCGAGCTGATAGCTCGATGTGACGGAGTCTCGGATTGCAAGACGGGGAGGAC
GAGTACCGCTGTGTCGGGTGGGTGGTCAGAATGCCGTGCTCCAGGTGTTTACAGCTGCTTC
GTGGAAGACCATGTGCTCCGATGACTGGAAGGTCACACGCAATGTTGCTGTGCCCAAC
TGGGTTTTCCCAAGCTATGTGAGTTCAGATAAACCAGAGTGAGCTCGCTGGAGGGGCAGTTTC
CGGGAGGAGTTTGTGTCCATCGATCACTCTTGCCAGATGACAAGGTGACTGCATTACACCA
CTCAGTATATGTGAGGGAGGGATGTGCTCTGGCCACGTGGTTACCTTGAGTGCACAGCTTC
GTGGTCATAGAAGGGGCTACAGCTCACGCATCGTGGGTGAAACATGTCTTGCTCTCGCAG
TGGCCTTGGCAGGCCAGCCTTCAGTTCCAGGGCTACCACTGTGCGGGGGCTCTGTCTATCAC
GCCCTGTGGATCATCACTGTGCACACTGTGTTTATGACTTTGACTCTTACCTCCCCAAGTCA
CCATCCAGGTGGGTCTAGTTCCCTGTGTGACAATCCAGCCCCATCCCACTGTGTGGAGAAG
ATTGCTTACCACAGCAAGTACAAGCCAAAGAGGCTGGGCAATGACATCGCCCTTATGAAGCT
GGCCGGGCCACTCAGTTCAATGAAATGATCCAGCCTGTGTGCTGCCCAACTCTGAAGAGA
ACTTCCCCGATGGAAGAGTGTGTGGACGTGAGGATGGGGGGCCACAGAGGATGGAGGTGAC
GCCTCCCTGTCTGAACACCGGGCCGTCCCTTTGATTTCACAAAGATCTGCAACACAG
GGACGTGTACGGTGGCATCATCTCCCCCTCCATGCTCTGCGCGGGCTACCTGACGGGTGGCG
TGACAGCTGCCAGGGGGAAGCGGGGGGCCCTGGTGTCTAAGAGAGGAGGCTGTGGAAG
TTAGTGGGAGCGACCACTTTGGCATCGGCTGCGCAGAGGTGAACAAGCCTGGGGTGTACAC
CCGTGTCACTCCTTCTGGACTGGATCCACGAGCAGATGGAGAGAGACCTAAAAACCTGAA
GAGGAAGGGGACAAGTAGCCACCTGAGTTCTGAGGTGATGAAGACAGCCCGATCCTCCCT
GGACTCCGTGTAGGAACCTGCACACGAGCAGACACCTTGGAGCTCTGAGTTCCGGCACCA
GTAGCAGGCCGAAAGAGGCACCCCTCCATCTGATTCCAGCAACAACCTTCAAGCTGCTTTT
GTTTTTTGTTTTTTTGGAGGTGGAGTCTCGCTCTGTTGCCAGGCTGGAGTGCAGTGGCGAAA
TCCCTGCTCACTGCAGCCTCCGCTTCCCTGGTTCAAGCGATTCTCTTGCTCAGCTTCCCA
GTAGCTGGGACCACAGGTGCCCGCCACCACACCCAACTAATTTTTGTATTTTAGTAGAGAC
AGGTTTTACCATGTGGCCAGGCTGCTCTCAAAACCCCTGACCTCAAATGATGTGCTGCTT
CAGCCTCCCAAGTGTCTGGGATTACAGGCATGGGCCACACGCTAGCCTCAGCTCCTTTC
TGATCTTCACTAAGAACAAAAGAACGACGAACTTGCAAGGGCGGCCTTTCCCACTGGTCCAT
CTGGTTTTCTCTCAGGGTCTTGCAAAATCCTGACGAGATAAGCAGTTATGTGACCTCAGG
TGCAAGCGACCAACAGCCTCAGAAAAGACGACCCAGCCAGAGTGCAGAACTGCAGTCT
ACTGCACGTTTTTCATCTCTAGGGACGAGAACCAACCCCTTTCTACTTCCAAGACTTAT
TTTCACATGTGGGGAGGTTAATCTAGGAATGACTCGTTTAAAGCCTATTTTCATGATTTCTT
TGTAGCATTTGGTGCTTGACGTATTATTGCTCTTGATTCCAAATAATATGTTTCTTCCCT
CATTTGCTGGCGTGTGCTGGGACTGGTGACGTGAATCAAAATCATCCACTGAAA

1007797.102401

FIGURE 28

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45234

><subunit 1 of 1, 453 aa, 1 stop

><MW: 49334, pI: 6.32, NX(S/T): 1

MGENDPPAVEAPFSFRSLFGLDDLKISPVAPDADAVAAQILSLLPLKFFPIIVIGIIALILA
LAIGLGIHFDCSGKYRCRSSFKCIELIARCDGVSDCKDGEDEYRCVRVGGQNAVLQVFTAAS
WKTMCSDDWKGHYANVACAQLGFPSYVSSDNLRVSSLEGQFREEFVSIHLLPDDKVTALHH
SVYVREGCASGHVVTIQCTACGHRRGYSSRIVGGNMSLLSQWPWQASLQFQGYHLCGGSVIT
PLWIITAAHCYVDLYLPKSWTIQVGLVSLLDNPAPSHLVEKIVYHSKYKPKRLGNDIALMKL
AGPLTFNEMIQPVCLPNSEENFPDGKVCWTSGWGATEDGGDASPVLNHAAVPLISNKICNHR
DVYGGIISPSMLCAGYLTGGVDSCQGDGGPLVCQERRLWKLVGATSFSGIGCAEVNKPQVYT
RVTSFLDWIHEQMERDLKT

Signal Peptide:

amino acids 1-20

Transmembrane domain:

amino acids 240-284

FIGURE 29

CCCACGCGTCCGTCTTAGTCCCGGGGCCAACTCGGACAGTTTGCTCATTATTATGCAACGGTCAAGGCTGGCTTGT
 GCTCAGAACCGCGCGCGCGCGCGCACACGACGACACACACGCGGGGAAACCTTTTAAATAAAGGCTAGAGA
 GCTCAGGCGCGCGCGCGCGCTCGCGGAGGGCTCGGAGCTGACTGCCGAGCGACAGGATCCCTCCGSGTCGGA
 CGCCCGGCGCGCGCTCGGCGCGCGCGTGGGATGGTGACGCGCTCGCGCGCGGCGCGGAGAGCTGCTGCACCTGAAG
 GCGCGCGACCAATGGCAGCGCGCGCGCTGCGCGTGTCCCGCGCGCGCGCTCTGCTCGCCCTTGCGCGGTGCTCT
 GCTCGCGCGCTGCGAGGCGCGAGGGGTGAGCTTATGGAACCAAGGAAGAGCTGATGAAAGTTTGTCAGTGCCTCTGT
 TCGGAGTGGGGACCTCTGGATCCGAGTGAAGAGCTTCGACTCCAAGAATCATTCAGAAAGTGCATGAATATTGCACT
 ACAACGGGAAGCAAGAACTGATCATAAATCTGGAAGAAATGAAGGTCTCATTGCCAGCATTTTCAAGGAAAC
 CCACATATCTGCAAGACGCTACTGATGTCTCCCTCGCTGAAATTACACGGGTCACTGTACTACCATGGACATGT
 ACGGGGATATTCTGATTACGAGCTCAGTCTCAGCACGTGTCTGGTCTCAGGGGACTTATTGTGTTTGAATAATGA
 AAGCTATGTCTTTAGAACCAATGAAAGTGCAACCAACAGATACAACTCTCCACGCGAAGAAGCTGAAAGCGT
 CCGGGGATCATGTGGATCACATCACACACACCAAACTCGCTGCAAGAATGTGTTTCCACCACCTCTCAGAC
 ATGGGCAAGAGGCATAAAAGAGAGACCTCAAGGCAACTAAGTATGTGGAGCTGGTGATCGTGGCAGACAAACG
 AGAGTTTCAAGAGCAAGGAAGATCTGGAAAAAGTTAAGCAGCGATTAAATAGAGATTGCTAATCACGTTGACAA
 GTTTTACAGACCACTGAACATTCGGATCGTGTGGTGAAGCGTGAAGTGTGGAATGACATGGACAAATGCTCTGT
 AAGTCAGGACCCATTCAACAGCTTCCATGAATTTCTGGACTGGAGGAAGATGAAGCTTCTACCTCGCAAAATCCCA
 TGACAATGCGAGCTTGTGAGTGGGTTTATTTCCAAGGACCACTCGGCATGGCCCCAATCATGAGCATGTG
 CACGGCAGACCACTCTGGGGGAATTTGTCATGGACCAATTGACAAATCCCTTGGTGACGGCGTGACCTTGGCACA
 TGAGCTGGGCGCAATTTTCGGGATGAATCATGACACACTGGACAGGGGCTGTAGCTGTCAAAATGGCGGTTGAGAA
 AGGAGGCTGCATCATGAACGCTTCCACCGGTAACCATTTCCCATGGTGTTCAGCAGTGCAGCAGGAAGGACTT
 GGAGACCAAGCTGGAGAAAGGATGGGGGTGTGCTGTTTAACTTGCAGGAGTCAAGGAGTCTTTTGGGGGCCA
 GAAGTGTGGGACAGATTTTGTGAAGAAGGAGAGGAGTGTGACTGTGGGGAGCCAGAGGAATGTATGAATCGCTG
 CTGCAATGCCAACCTGTACCTGTGAAGCCGGAGCGCTGTGTGCGCACCTGGAGCTGTGCTGTGAAGACTGCCAGCT
 GAAAGCTCGAGGAACAGGTGTCAGGACTCCAGCAACTCTGTGACCTCCAGACTTCTGCAACAGGGCGACGCC
 TCACTGCCAGCCAAATGTGTACTGTGACAGTGGGCACTCATGTGAGATGTGGACGGCTACTGCTCAATGGCAT
 CTGCGCACTCAAGACAGCAGCTGTCTCACGCTCTGGGACAGGTGCTCAACCTGGGATCTGCTTTTGA
 GAGACTCAATTTCTGCGAGTGTATCTTATGGCACTGTGGCAAGTCTCGAAGAGTTCTTTGCGAAATGCCAGAT
 GACAGATGCTAAATGTGAAAAATCCAGTGTCAAGGAGGTGCCAGCGGCAGTCAATGTACCAATGCCCTTTC
 CATAGAAACAAACATCCCTCTGACGCAAGGAGGCGGATTTCTGTGCGCGGGGAGCCACGTGTACTTGGCGCATGA
 CATGCGCGGACCCAGGCTTTGTGCTTGCAGGCACAAGGTGTGCAGATGGAATAATCTGCTGTAATCGTCAATGTCA
 AAATATTAGTGTCTTTGGGGTTCAAGAGTGTGCAATGCAAGTGCACGCGCAGAGGGGTGTGCAACACAGGGAAGAA
 CTGCCACTTGGGAGGCCCACTGGGCACTCCCTCTGTGTACAGTTTGGCTTTGGAGGAAGCAGACAGACGCGGCC
 CATCGGCAAGCAGAAAGCAAGGCGAGGAAGCTGCAGAGTCCAAAGGAGCGCGGCGCGGAGGAGCCCGTGGG
 ATCGCAGGAGCATGCGTCTACTGCTCACTGACACTCATCTAGAGCCCTCCCATGACATGGAGACCGTGACCACTG
 CTGCTGCGAGGAGGTTCAGCGTCCCCAAGGCTCTGTGACTGGCAGACTTACTCTGTGGCTTTGGCATCGTT
 TCCATGACACAGACACAACTAGTTCTCGGGCTCAGGAGGGGAAGTCCAGCCTACCAAGGACGCTTGCAGAAA
 CAGTGCAGGAAGGGCAGCAGCTTCTGTTGAGCTTCTGCTAAAAACATGGACGCTTCACTGCTGCTGCTCTGAG
 AGAGTAGCAGGTTTACCATCTTGGCAGGCGCCAGCCCTGCAGCAAGGAGGAAGAGGACTCAAAAGTCTGGCCCTTC
 ACTGAGCTCCACAGCAGTGGGGGAGAAAGCAAGGTTGGGCGCACTGTCCCTTTCCCACTGACCACTCAGCCT
 TGGCAGCCCTGATGACTGGTCTCTGGCTGCAACTTAATGCTCTGATATGGCTTTTAGCATTTTATATAGAAAT
 AGCAGGGTTTTAGTTTTTAAATTTATCAGAGACCTGCCACCATTCATCTCCATCAAGCAAACTGAATGGCAA
 TGAAACAACTTGAGAAAGAGGTAGGAGAAAGGGCGGTGAATCTGCTCTTGTCTGTGACATCGCTGACGACCA
 AGTACTCAGTTTGAAGGTTTGCAAGAGCCAGGGAACCAACAGAGTCAACCACTCTATTAAACAGTGAAGAA
 TGTAAAAAGTGAACAAATGTAAAGGCTAACTCATCCCCGTGGCCATTACTGCATAAAATAGAGTGCAATTT
 GAAAT

10017391.1000101

FIGURE 30

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49624

><subunit 1 of 1, 735 aa, 1 stop

><MW: 80177, pI: 7.08, NX(S/T): 5

MAARPLPVSPARALLALAGALLAPCEARGVSLWNQGRADEVVVSASVRSGDLWIPVKSFDSK
NHPEVLNIRLQRESKELIINLERNEGLIASSFTETHYLQDGTDVSLARNYTGHCYYHGHVRG
YSDSAVSLSTCSGLRGLIVFENESYVLEPMKSATNRYKLFPAKKLKSVRGSCGSHHNTPNLA
AKNVFPPPSQTWARRHKRETLKATKYVELVIVADNREFQRQGDLEKVKQRLIEIANHVDKF
YRPLNIRIVLVGVEVWNDMKCSVSQDPFTSLHEFLDWRKMKLLPRKSHDNAQLVSGVYFQG
TTIGMAPIMSMCTADQSGGIWMDHSDNPLGAAVTLAHELGHNFGMNHDTLDRGCSCQMAVEK
GGCIMNASTGYPPPMVFSSCSRKDLTSLEKGMGVCLFNLPEVRESFGGQKCGNRPFVEEGEE
CDCGEPEECMNRCNATTCTLKPDVAHGLCCEDCQLKPAGTACRDSNSCDLPEFCTGAS
PHCPANVYLHDGHSQDVGICYNGICQTHEQQCVTLWGPAGKAPAGICFERVNSAGDPYGN
CGKVSKSFFAKCEMRDAKCGKIQCGGASRPVIGTNAVSIETNIPLQQGGRIILCRGTHVYLG
DDMPDPGLVLAGTKCADGKICLNRQCQNISVFGVHECAMQCHGRGVCNNRNKNCHEAHWAPP
FCDKFGFGGSTDSGPIRQAEARQEAAESNRERGQGQEPVGSQEHASTASLTLI

Signal peptide:

amino acids 1-28

TCCCAAGGCTTCTTGGATGGCAGATGATTNTGGGGTTTTGCATTGTTTCCCTGACAAACGAAA
ACAAAACAGTTTTGGGGGTTTCAGGAGGGGAANTCCAGCCTACCCAGGAAGTTTGAGAAAAC
GTGCAAGGGAAGGCAGAGANTTCTCGTTGAGNTTTTTGNTAAACATGGACATGNTTCAGTG
CTGCTCNTGAGAGAGTAGCAGGTTACCACTTTTGGCAGGCCCCAGCCTGCAGCAAGGAGGA
AGAGGACTCAAAAGTTTGGCCTTTCACTGAGCCTCCACAGCAGTGGGGGAGAAGCAAGGGTT
GGGCCCAAGTAGTCCCCTTTCCCAGTAGCACCTCAGCCTTGGCAGCCCTGATAACTGGTNTNT
GGCTCCANTTAATGCNTGATATGCTTTTTAGCACTTTATTATTAATAAGAAATAGCAGGGTTTT
AGTTTTTAATTTATCAGAGACCTGCCACCACTTCCATNTCCATCCAAG

FIGURE 32

CATCCTGCAACATGGTGAAACCACGCCTGGCTAATTTTGTGTATTTTGGTAGAGATGGGA
TTTCACCGTGTTAGCCAGGATTGTCTCAATCTGACCTCATGATCTGCCCCCTCGGCCCTCC
AAGTGCTGGGATTACAGCGGAGTGCAACCACACCCGGCCACAACTTTTAAAGAGTTAAT
GAAACCATACCTTTTACATTTTAAATGACAGGAAAATGCTCACAAATAATTGTTAAACCCAAA
TTCTGGGATACAAAAGTACAATCTTTACTGTGTAAATACATGTATATGTACTATATGAAAATA
TACCAAAATATCAATAATACTTATCTCTGGGTAAAAACCTCTTCTCATACCCGTGTGCTAACAA
CTTTTAACAAAAAATTTGCATCACTTTTAAAGATCAAGAAAAATTTCTGAAGGTCATATGGG
ACAGAAAAAAAACCAAGGGAAAAATCACGCCACTTTGGGAAAAAAGATTGAAATCTGCCT
TTTTATAGATTGTAAATTAATAAGGTCAGGCTTTCTAAGCAACTTAATGTTTTGTTTCGA
AACAAAGTACTTGTCTGGATGTAGGAGGAAAGGAGTGATGTCCTGCCATTATGATGCCCC
TTGAATATAAGACCTACTTGTCTATCTCCCTGCACCAGCCAGGAGCCACCCATCTCCAGC
ACACTGAGCAGCAAGCTGGACACACGGCACCTGATCCAAATGGGTAAGGGGATGGTGGCGA
TGCTCATTCTGGGTCTGCTACTTCTGGCGCTGCTCCTACCCGTGCAGGTTTCTTCATTTGTT
CCTTTAACCAGTATGCCGGAAGCTACTGCAGCCGAAACCACAAAGCCCTCCAACAGTGCCCT
ACAGCTACAGCCGGTCTCCTTGTGGTCTTGCTTGCCCTTCTACATCTCTACCATTAAAGAGG
CAGGTCAAGAAACAGCTACAGTTCTCCAACCATACACTAAAACCGAATCCAAATGGTGCCT
AGAAGTTCAATGTGGCAAGGAAAAAACAGGTCTTCATCAAACTACTAATTTCACTCCTT
ATTAACAGAGAAACGCTTGAGAGTCTCAAACCTGGACTGGTTTAAAGAGCATCTGAAGGATTT
GACTAGATGATAAATGCCTGTACTCCAGTACTTTGGGAGGCTAGGCCGGCGGATCACCTG
AGGTGAGGATTTGAGACTAACCTGGCCAAAATGGTGAAACCCCATCTGTACTAAAAATACA
AATATTGACTGGCGTGGTGGTGAGTGCCGTGTGATCCAGCTACTCAGGTGGCTGAAGCAGG
ACAATCACTTGAACCTCAGGAGGCAGAGGTTGCAGTGAGCTGAGATCGCGCTACTGCACTCTA
GCCTAGCCTGGGCAACAGAGTGAGACTTCGTCTCAAAAAAAAAAAGCCAAGTGCACTGGCT
CACGCCGTGAATCCCGGCCTTTGGGAGGCCGAGGTGGGCGGATCACGAGGTGAGGAGATCA
AGACCATCTGGCTAATAACAGTGAAACCTGTCTCTACTAAAAATACAAAAATAGCCGGG
GATGGTGGCAGGCACCTGGAGTCCAGCTACTCGGAGGCTGAGGCAGGAGAATAGCGTGAA
CTCAGGAGGCCGAGCTTGCAGTGAGCCGAGATTGCGCTACTGCACTCCAGCCTGGGCGCAG
CGCGAGACTCCGTCTCAAAAAAAAAAAAAAAAAAAAAA

FIGURE 33

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48309

><subunit 1 of 1, 67 aa, 1 stop

><MW: 6981, pI: 7.47, NX(S/T): 0

MGKGMVAMLILGLLLALLLPVQVSSFVPLTSMPEATAAETTKPSNSALQPTAGLLVVLLAL
LHLVH

Signal peptide:

amino acids 15-27

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

[illegible][illegible]

FIGURE 35

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA46776

><subunit 1 of 1, 432 aa, 1 stop

><MW: 47629, pI: 5.90, NX(S/T): 0

MPARPGRLPLLLARPAALTALLLLLLGHGGGGRWGARAQEAAAAAADGPPAADGEDGQDPHS
KHL YTADMFTHGIQSAAHFVMFFAPWCGHCQRLQPTWNDLGDKYNSMEDAKVYVAKVDCTAH
SDVCSAQGVRGYP TLKLFKPGQEAVKYQGPRDFQTL ENWMLQTLNEEPVTPEPEVEPPSAPE
LKQGLYEL SASNFELHVAQGDHFIKFFAPWCGHCKALAPTWEQLALGLEHSETVKIGKVDCT
QHYELCSGNQVRGYPTLLWFRDGKKVDQYKGRDLES LREYVESQLQRTETGATETVTPSEA
PVLAAEPEADKGTVLALTENNFD DTTIAEGITFIKFYAPWCGHCKTLAPTWEELSKKEFPGLA
GVKIAEVDCTAERNICSKYSVRGYPTLLLPFGGKKVSEHSGGRDLDSLHRFVLSQAKDEL

Signal sequence:

amino acids 1-32

001701-102401

FIGURE 36

CTTTTCTGAGGAACCACAGCAATGAATGGCTTTGTCATCCTTGCTTCGAAGAAACCAATTTAT
CCTCCTGGTACTATTTCTTTTGCAAATTCAGAGTCTGGGTCTGGATATTGATAGCCGTCCTA
CCGCTGAAGTCTGTGCCACACACAATTTACCAGGACCCAAAGGAGATGATGGTGAAAAA
GGAGATCCAGGAGAAGAGGGAAAGCATGGCAAAGTGGGACGCATGGGGCCGAAAGGAATTAA
AGGAGAACTGGGTGATATGGGAGATCAGGGCAATATTGGCAAGACTGGGCCCATTTGGGAAGA
AGGGTGACAAAGGGGAAAAAGGTTTGCTTGGAATACCTGGAGAAAAAGGCAAAGCAGGTACT
GTCTGTGATTGTGGAAGATACCGGAAATTTGTTGGACAACCTGGATATTAGTATTGCTCGGCT
CAAGACATCTATGAAGTTTGTCAAGAATGTGATAGCAGGGATTAGGGAAACTGAAGAGAAAT
TCTACTACATCGTGCAGGAAGAGAAGAACTACAGGGAATCCCTAACCCACTGCAGGATTCGG
GGTGAATGCTAGCCATGCCCAAGGATGAAGCTGCCAACACACTCATCGCTGACTATGTTGC
CAAGAGTGGCTTCTTTCCGGTGTTCATTGGCGTGAATGACCTTGAAAGGGAGGGACAGTACA
TGTCACAGACAACACTCCACTGCAGAACTATAGCAACTGGAATGAGGGGAACCCAGCGAC
CCCTATGGTCATGAGGACTGTGTGGAGATGCTGAGCTCTGGCAGATGGAATGACACAGAGTG
CCATCTTACCATGTACTTTGTCTGTGAGTTCATCAAGAAGAAAAAGTAACTTCCCTCATCCT
ACGTATTGCTATTTTCTGTGACCGTCATTACAGTTATTGTTATCCATCCTTTTTTCTCTG
ATTGTACTACATTTGATCTGAGTCAACATAGCTAGAAAATGCTAAACTGAGGTATGGAGCCT
CCATCATCAAAAAAAAAAAAAA

100191-10401

FIGURE 37

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50980

><subunit 1 of 1, 277 aa, 1 stop

><MW: 30645, pI: 7.47, NX(S/T): 2

MNGFASLLRRNQFILLVLFLLQISLGLDIDSRPTAEVCATHTISPFGKGDGEGKDPGEEG
KHGKVGRMGPKGIKGE LGDMGDQGNIGKTGPIGKKGDKGEKGLLGIPGEKGKAGTVCDGCRY
RKFBVGQLDISIARLKTSMKFVKNVIAIGIRETEEFYIIVQEEKNYRESLTHCRIRGGM LAMP
KDEAANTLIADYVAKSGFFRVFIGVNDLREGEQYMSTDNTPLQNYSNWNEGEPSDPYGHEDC
VEMLSGRWNDECHLTMYFVCEFIKKKK

Signal peptide:

amino acids 1-25

FIGURE 38

GGTTCTATCGATTGCAATTCGGCCACACTGGCCGGATCCTCTAGAGATCCCTCGACCTCGAC
CCACGCGTCCGCTGCTCTCCGCCGTGTGGAGTGGTGGGGCCCTGGTGGGAATGGCGTGT
GCCAGCGCACGCGCGCTCCCTGGAAGGAGAAGTCTCAGCTAGAACGAGCGGCCCTAGGTTT
CGGAAGGGAGGATCAGGGATGTTTTCGAGCGCGCTGGAACGAGCGGTGCCGATAGAGGAAGC
GGGCTCCATGGTGCCCTCTGTGCTGCCCTGCTGCTGTTGCTACCGCTGCTGCTGCTGA
AGCTACACCTCTGGCCGAGTTGCGCTGGCTTCCGGCGGACTTGGCCTTTGCGGTGCGAGCT
CTGTGCTGCAAAAGGGCTCTTCGAGCTCGCGCCTGGCCGCGTGGCTGCGCCGACCCGGAAGG
TCCCAGGGGGGCTGCAAGCTGGCTGGCGCTCGCGGAAGTGGCCAGCAGCGCCCGCGC
ACACCTTTCTCATTACGGCTCGCGCGCTTTAGCTACTCAGAGCGGAGCGCGAGAGTAAC
AGGGCTGCACGCGCCTTCTACGTGCGCTAGGCTGGACTGGGGACCCGACGCGCGCGACAG
CGGCAGGGGAGCGCTGGAGAAGGCGAGCGGCGAGCGCGGAGCCGAGAGATGCAAGCGCCG
GAAGCGCGCGGAGTTTGGCGGAGGGGACGGTGCCCGCAGAGTGGAGGAGCCCGCCCT
CTGTCACTGGAGCAACTGTGGCGTGTCTCTCCCGCTGGCCAGAGTTTCTGTGGCTCTG
GTTCCGGCTGGCCAAGGCCGCGCTGCGCACTGCTTTGTGCCACCTGCGCTGCGCGGGCC
CCCTCTGCTGCACTGCCTCGCAGCTGCGCGCGCGCGCTGGTCTGGCGCCAGAGTTTCTG
GAGTCCCTGGAGCCGACCTTGCCTCCCGCTGAGAGCCATGGGGCTCCACCTGTGGGTGCAAG
CCCAGGAACCCACCTGCTGGAATTAGCGATTGCTGGCTGAAGTGTCCGCTGAAGTGGATG
GGCAGTGGCAGGATACCTCTCTTCCCGCAGAGCATAACAGACAGCTGCCGTGTACATTTT
ACCTCTGGCACCACGGGCCCTCCCAAGGCTGCTCGGATCAGTCATCTGAAGATCTCCTGAATG
CCAGGGCTTCTATCAGCTGTGTGGTGTCCACCAGGAAGATGTGATCTACCTCGCCCTCCAC
TCTACCACATGTCCGGTTCCTGTGCGGATCTGCGGTGTCATGGGCATTTGGGGCCACAGTG
GTGCTGAATCCAGTTTCTCGGTGGTCACTTCTGGGAAGATTGCCAGCAGCACAGGGTGAC
GGTGTTCAGTACATTGGGGAGCTGTGCCGATACCTTGTCAACAGCCCGCGAGCAAGGAGCAG
AACGTGGCCATAAGGTCCGGCTGGCAGTGGGCAGCGGGCTGCGCCAGATACCTGGGAGCGT
TTTGTGCGCGCTTCGGGCCCTGCAGGTGTCTGGAGACATATGGACTGACAGAGGGCAACGT
GGCCACCATCAACTACACAGGACAGCGGGCGCTGTGGGGCGTCTTCTTGCTTTACAAGC
ATATCTTCCCTTCTCCTTGATTGCTATGATGTCAACAGGAGGACCAATTCGGGACCCC
CAGGGGCACTGTATGGCCACATCTCCAGGTGAGCCAGGGCTGCTGGTGGCCCGGTAAAGCA
GCAGTCCCATTTCTTGGGCTATGCTGGCGGGCCAGAGCTGGCCAGGGGAAGTTGCTAAAGG
ATGTCTTCCGGCTGGGGATGTTTTCTTCAACACTGGGACCTGCTGGTCTGCGATGACCAA
GGTTTTCTCCGCTTCCATGATCGTACTGGAGACACCTTCAGGTGGAAGGGGAGAATGTGGC
CACAAACGAGGTGGCAGAGGTCTTCGAGGCCCTAGATTTTCTTCAGGAGGTGAACGCTCTATG
GAGTCACTGTGCCAGGGCATGAAGGCAAGGCTGGAATGGCAGCCCTAGTTTCTGCGTCCCCC
CACGCTTTGGACCTTATGCACTCTACACCCACGTGTCTGAGAAGTTGCCACCTTATGCCCG
GCCCGGATTCCTCAGGTCTCAGGAGTCTTTGGCCACCAAGAGACCTTCAAACAGGAGAAAG
TTCCGATGGCAATGAGGGCTTCGACCCGACCCCTGTCTGACCCACTGTACCCCTTCTGAGC
CAGGCTGTAGGTGCTTACCTGCCCTCACACTGCCCGGTACAGCGCCTCTGCGCAGGAAA
CCTTCGAATCTGAGAAGTTCCACACCTGAGGCACCTGAGAGAGGAACTCTGTGGGGTGGGG
CCGTGCGAGGTGACTGGGTGTGAGGGATCTTTCTATACCAAGACTGCGGTCACTATTTT
GTAATAAATGTGGCTGAGCTGATCCAGCTGTCTCTGACCTAAAAAAGGGGATTAAGC
ATAAAGGCGCGCGGACTCTAGAGTGCACCTGCACTAGGAGTAACAGGGTAAAGC
TTGGCCGCATGGCCCACTTGTATTGACG

001791.102401

FIGURE 39

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50913
><subunit 1 of 1, 730 aa, 1 stop
><MW: 78644, pI: 7.65, NX(S/T): 2
MGVCQRTAPWKEKSQLERAALGFRKGGSGMFASGWNQTVPIEEAGSMAALLLLPLLLLLPL
LLLKLHLWPQLRWLPADLAFVRLALCCKRALRARALAAAAADPEGPEGGCSLAWRLAELAQQ
RAAHTFLIHGSRFRFSYSEAERESNRAARAFLRALGWDWGPDDGSDGEGSAGEGERAAPGAGD
AAAGSGAEFAGGDGAARGGGAAPLSPGATVALLLPAGPEFLWLWFGGLAKAGLRATFVPTAL
RRGPLLHCLRSCGARALVLAPEFLESLEPDLPALRAMGLHLWAAGPGTHPAGISDLLAEVSA
EVDGVPVPGYLSPPQSITDTCLYIFTSGTTGLPKAARISHLKILQCQGFYQLCGVHQEDVIYL
ALPLYHMSGSLLGIVGCMGIGATVVLLKSKFSAGQFWEDCQQHRVTVFQYIGELCRYLVNQPP
SKAERGHKVRRLAVGSGRLPDTWERFVRRFGPLQVLETYGLTEGNVATINYTGQRGAVGRASW
LYKHIFPFSLIRYDVTITGEPIRDPPQGHCMATSPGEPGLLVAPVVSQQSPFLGYAGGPFLAQGK
LLKDVFRPGDVFFNNTGDLVLCDDQGFLRFHDDRTGDTFRWKGENVATTEVAEVFEALDFLQEV
NVYGVTVPGHEGRAGMAALVLRPPHALDLMQLYTHVSENLPFYARPRFLRLQESLATTETFK
QQKVRMANEGFDPSTLSDPLYVLDQAVGAYLPLTTARYSALLAGNLRI

Type II transmembrane domain:

amino acids 45-65

Other transmembrane domain:

amino acids 379-398

cAMP- and cGMP-dependent protein kinase phosphorylation site
starting at amino acid 136

CUB domain protein motif

amino acids 254-261

putative AMP-binding domain signature

amino acids 332-343

N-glycosylation sites

amino acids 37-40 and 483-486

FIGURE 40

CCTGTGTTAAGCTGAGGTTTCCCCTAGATCTCGTATATCCCCAACACATACCTCCACGCACA
CACATCCCCAAGAACTCGAGCTCACACCAACAGACACACGCGCGCATACACACTCGCTCTC
GCTTGTGCTCATCTCCCTCCCGGGGAGCGCGCGCGCTCCCACTTTGCCGCACACTCCGGC
GAGCCGAGCCCGCAGCGCTCCAGGATTCTGCGGCTCGGAACTCGGATTGCGAGCTTGAAGCCC
CCATGGTGGTTTTTTAAACACTTCTTTTCTCTCTTCTCGTTTTGATTGCACCGTTTCCA
TCTGGGGCTAGAGGAGCAAGGCAGCAGCCTTCCCAGCCAGCCCTTGTGGCTTGCATCGT
CCATCTGGCTTATAAAGTTTGTGAGCGCAGTCCAGAGGGCTGCGCTGCTCGTCCCCTCGG
CTGGCAGAAGGGGGTGACGCTGGGCAGCGCGAGGAGCGCGCGCTGCTCTGGCGGGCTTT
CGGCTTGAGGGGCAAGGTGAAGAGCGCACCCGGCCGTGGGGTTTACCAGAGCTGGATTGTATG
TTGCACCATGCTTCTTGGATCGGGGCTGTGATTCTTCCCCTCTTGGGGCTGCTGCTCTCCC
TCCC CGCGGGGCGGATGTGAAGGCTCGGAGCTCGCGAGAGGTCCGCGCAGGCGTACGGTGCC
AAGGAGTTACGCTTGGCGACATCCCCTACCAGGAGATCGCAGGGGAACACTTAAGAATCTG
TCCTCAGGAATATACATGCTGCACCCACAGAAATGGAAGACAAGTTAAGCCAACAAAGCAAAC
TCGAATTTGAAAACCTTGTGGAAGAGACAAGCCATTTGTGCGCACCACTTTGTGTCCAGG
CATAAGAAATTTGACAAATTTTCCGAGAGCTCCTGGAGAATCGAGAAAAGTCACTAAATGA
TATGTTTTGTACGGACCTATGGCATGCTGTACATGCAAGAATTAGAAGTCTTCCAGGACCTCT
TCACAGAGCTGAAAAGGTACTACACTGGGGGTAATGTGAATCTGGAGGAAATGCTCAATGAC
TTTTGGGCTCGGCTCCTGGAACGGATGTTTCAGCTGATAAACCCCTCAGTATCACTTCAGTGA
AGACTACCTGGAATGTGTGAGCAATACACTGACCAGCTCAAGCCATTTGGAGACGTGCCCC
GGAAACTGAAGATTCAGGTTACCCGCGCCTTATTGCTGCCAGGACCTTTGTCCAGGGGGTG
ACTGTGGGCAGAGAAGTTGCAAACCGAGTTTCCAAGGTGAGCCCAACCCAGGGTGTATCCG
TGCCCTCATGAAGATGCTGTACTGCCATACATGTGCGGGGCTTCCCCTGTGAGGCCCTGCA
ACAACTACTGTCTCAACGTTCATGAAGGCTGTCTGGCAAATCAGGCTGACCTCGACACAGAG
TGGAATCTGTTTTATAGATGCAATGCTCTTGGTGGCAGAGCGACTGGAGGGGCCATTCAACAT
TGAGTCGGTCATGGACCCGATAGATGTCAAGATTTCTGAAGCCATTATGAACATGCAAGAAA
ACAGCATGCAGGTGTCTGCAAAGGTTCTTTCAGGGATGTGGTCAGCCCAAACCTGTCTCAGCC
CTCAGATCTGCCCGCTCAGCTCCTGAAAATTTTAATACAGTTTCAGGCCCTACAATCCTGA
GGAAAGACCAACAACCTGCTGCGAGGCAACAGCTTGGACCGGCTGGTCACAGACATAAAGAGA
AATTGAAGCTCTCTAAAAAGGTCTGGTCAGCATTACCTTACACTATCTGCAAGGACGAGAGC
GTGACAGCGGGCAGCTCCAACGAGGAGGAATGCTGGAACGGGCACAGCAAAGCCAGATACTT
GCCTGAGATCATGAATGATGGGCTCACCAACCAGATCAACAATCCCGAGGTGGATGTGGACA
TCACTCGGCCTGACACTTTCATCAGACAGCAGATTATGGCTCTCCGTGTGATGACCAACAAA
CTAAAAACGCTACATCAATGGCAATGATGTCAATTTCCAGGACAAGTGATGAATCCAGTGG
CTCAGGGAGTGGCAGTGGGTGCAATGGATGACGTGTGTCACCGAGTTTGTAGTTTGTACCA
CAGAGCCCCCGCAGTGGATCCCGACCGGAGAGAGGTGGACTCTTCTGCAGCCACGCGTGGC
CACTCCCTGCTCTCCTGCTCTCTCACCTGCATTGTCCTGGCACTGCAAGACTGTGCAGATG
ATCTTGGGTTTTTGGTCAGATGAACTGCATTTTAGCTATCTGAATGGCCAACCTCACTTCTT
TTCTTACACTCTTGGACAATGGACCATGCCACAAAACCTTACCGTTTTCTATGAGAAGAGAG
CAGTAACTGCAATCTGCTCCCTTTTGTTTTCCAAAGAGTACCGGGTGCCAGACTGAACCTG
CTTCTCTTTCTTCTTCACTATCTGTGGGACCTTGTATTATTCTAGAGAGAAATCTTACTCAA
ATTTTTCTGACCGAGAGATTTTCTTACCTTCATTGTCTTTATGCTGCAAGAGTAAAGGAAT
CTCAGTTGTGAGGGTTTTTTTTTCTCATTATAAAT

FIGURE 41

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50914

><subunit 1 of 1, 555 aa, 1 stop

><MW: 62736, pI: 5.36, NX(S/T): 0

MPSWIGAVILPLLGLLLSLPAGADVKA RSCGEVRQAYGAKGFS LADIPYQEIAGEHLRICPQ
EYTCCTTEMEDKLSQQSKLEFENLVEETSHFVRTTFVSRHKKFDEFFRELL ENAEKSLNDMF
VRTYGMLYMQNSEVFQDLFTELKRYYTGGNVNLEEMLNDFWARLLERM FQLINPQYHFS EDY
LECVSKYTDQLKPFQGDVPRKLKIQVTRAFIAARTFVQGLTVGREVANRVSKVSP TPGCIRAL
MKMLYCPYCRGLPTVRPCNNYCLNVMKGCLANQADLDTEWNLFIDAMLLVAERLEGP FNIES
VMDPIDVKISEAIMNMQENSMQVS AKVFQGCQPKPAPALRSARSAPENFNTRFRPYNPEER
PTTAAGTSLDRLVTDIKEKLKLSKKVWSALPYTICKDESVTAGTSNEEECWNGH SKARYLPE
IMNDGLTNQINNPEVDVDITRPDTFIRQQIMALRVMTNKLKNAYNGNDVNFQDTSDESSGSG
SGSGCMDVDCPTEFEFVTT EAPAVDPDRREVDSSAAQRGHSLSWSLTCIVLALQRLCR

Signal peptide:

amino acids 1-23

FIGURE 42A

CGGACGCGTGGGCGGACGCGTGGGCAAAAGAACTCGGAGTGCCAAAGCTAAATAAGTTAGCTGAGAAAAACGCACG
 CAGTTTGACGCGCTTCGCGCCGGTGCCTCACTACGCAAGACCAAGCGGGCTCCGCGCGGACCGCGCGCGGGG
 TAGGGACCCGGCTTTGGCCTTCAGGCTCCCTAGCAGCGGGGAAAAGGAATGCTGCCCCGAGTTTCTGCGGAGGT
 GGAGGGAGATCAGGAACCGGCTTCTTCTCACTTCGCCGCTGGTGAAGTGTGGGGAGATTGGCAACCGCTAGG
 AAAGGACTGGGGAAAATAGCCCTGGGAAAGTGGAGAAGGTGATCAGGAGGCCGGTCCACTACGGCAGTTTATCTG
 TCTGATCAGAGCCGACGCGCAGCGCTCCACTTCGCAGTTCTTTCCAGTGTGGGGACCGCAGGACAGACGCGCCGA
 TCCCGCCGCCCTCCGTACCGACCTCCAGGAGAGTCAAGCTCGCTCCCCAACGTCAGGGCGCTCTGGCCACGA
 AAAGTTCCTGTCCACTGTGATTCTCAATTCCTTGCTTGGTTTCTTCAGAGAAGTTTGGTGGAGATATTAT
 ACTTTTTCTTTTTTTTTCTTGGTGGAAAGTGTCTTAGGGAGGGGGAGGAGGAGGAGAAGTGAATGTGC
 TGGAGAAGAGCGAGCCCTCTTGTCTTCGGAGTCCCATCCATTAGGCCATCACTTCTGGAAGATTAAAGTTGT
 CGGACATGGTGACAGCTGAGAGGAGAGAGGATTCTTGCCAGGTGGAGAGTCTTACCGTCTGTGGGTGCATG
 TGTGCGCCCGCAGCGCGCGGGCGGTGTTCTCCGCTGGAGTCTCACTGGGACCTGAGTGAATGGCTCCCA
 GGGGCTGTGCGGGCATCCGCCTCCGCTTCTCCACAGGCGTGTGTCTGTCTGGAAGATGCTAGCAATGGGG
 CGCTGGCAGGATTCTGGATCTCTGCCTCTCACTTATGTTACCTGTCTGGGGCAGGCGCTTAGAAGAGGAGG
 AAGAAGGGGCTTACTAGCTCAAGCTGGAGAGAACTAGAGCCAGCACACTTCCACTCCAGCCCTCACTCA
 TTTTTCATCCTAGCGGATGATCAGGATTTAGAGATGTGGTTACACCGATCTGAGATTAAACACCTACTCTTG
 ACAAGCTCGTGCCGAAGGAGTTAAACTGGAGAATCTATGTCCAGCCTATTGACACCATCCAGGAGTCAGT
 TTATTACTGGAAGTATCAGATACACACCGGACTTCAACATTCTATCATAGACCTACCCAAACCACTGTTTAC
 CTCTGGCAATGCCACCTCACTCAGAACTGAAGGAGGTGGATTTCAACGCATATGGTCGGAAGTGGCAT
 TGGGTTTTAAACAGAAAAGATGATGCCACCAAGAGAGGATTTGATACCTTTTGGTTCCTTTTGGGAAGTG
 GGGATTACTATACACTACAAATGTGACAGTCTGGGATGTGTGGCTATGACTTGTATGAAAACGACAAATGCTG
 CCTGGGACTATGACAAATGGCATATACTCCACACAGATGTACACTCAGAGAGTACAGCAAACTCTAGCTTCCATA
 ACCCCACAAGCCTATATTTTATATACTGCCTATCAAGCTGTTCACTCACCACCTGCAAGCTCCTGGCAGGTATT
 TCGAACACTACCAGTCCATTATCAACATAAACAGGAGAAGATATGCCATGCTTCTCCTGTAGATGAAGCAA
 TCAACAACGTGACATTGGCTCTAAAGACTTATGTTTCTATAACAACAGCATTATCACTTACTCTTCAGATAATG
 GTGGCAGCCTACGGCAGGAGGGAGTAACTGGCTCTCAGAGGTAGCAAGGAACATATTGGGAAGGAGGATCC
 GGGCTGATAGGCTTTGTGCATAGCCCACTCTGAAAAACAAGGGAACAGTGTGTAAGGAACCTGTGACATCACTG
 ACTGTTACCCCACTCTCACTTCTCACTGCTGAAGGACAGATTGATGAGGACATTCACTAGATGGCTATGATCT
 GGGAGACATAAGTGAGGCTCTGCTCACCCGAGTAGATATTTGCATAAACATTGACCCCTATACCAAGGC
 AAAAAATGGCTCTTGGGCAGCAGGCTATGGGATCTGGAACACTGCAATCCAGTCAGCCATCAGAGTGCAGCACTG
 GAAATTGCTTACAGGAAATCTTGGCTACAGCGACTGGGTCCCCCTCAGTCTTTACAGCAACTGGGACCGAACCG
 GTGGCAATGAACGGATCACTTGTCACTGGCAAAAGTGTATGGCTTTTCAACATCACAGCCGACCCATATGA
 GAGGTGGACCTATCTAAACAGGTATCAGGAATCGTGAAGAAGCTCCTACGGAGGCTCTCACAGTTCAACAAAAC
 TGCAAGTCCGGTCAAGTATCCCCCAAGACCCAGAAGTAACCTAGGCTCAATGGAGGGGTCTGGGGACCATG
 GTATAAGAGGAAACCAAGAAAAGAGGCCAAGCAAAAATCAGGCTGAGAAAAGCAAAAGAAAAGCAAAAAA
 GAAGAAGAAACAGCAGAAGCAGTCTCAGGTAACACAGCAAAATTTGGCTCGATAATATCGCTGGCCTAAGCGTCA
 GGCTTGTCTTTCATGCTGTGCCACTCCAGAGACTTCTGCCACCTTGGCCGCCCACTGCAAACTGTCTCTGCAAG
 CCAAGTGTCTACTCTTGCAGAGCCACTTAGAGAGAGTGGAGATGTTATTTCTCTGCTCCTTTAGAAAACGTG
 GTGAGTCTGAGTTCACCTGCTGTGCTTCACTCACTGACCAACACTGCTTTGAATTATAGGAGGAGAACATA
 ACCTACCTATCGCAAGCATGCTAATTTGATGGAAGTTACAGGGTAGCATGATTAACACTACCTTTGATAAATTAC

[illegible]

FIGURE 43

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48296
><subunit 1 of 1, 515 aa, 1 stop
><MW: 56885, pI: 6.49, NX(S/T): 5
MAPRGCAGHPPPPSPQACVCPGKMLAMGALAGFWILCLLTGYLSWGQALEEEEEGALLAQA
GEKLEPSTTSTSQPHLIFILADDQGFRDVGYPHGSEIKTPTLDKLAEEGVKLENYVQPICTP
SRSQFITGKYQIHTGLQHSIIIRPTQPNCLPLDNATLPQKLKEVGYSTHVMGKWHLGFNKEC
MPTRRGFDTFFGSLLGSGDYIYTHYKCDSPGMCGYDLYENDNAAWDYDNGIYSTQMYTQRVQQ
ILASHNPTKPIFLYTAYQAVHSPLOAPGRYFEHYRSIININRRRYAAMLSCLEAINNVTLA
LKTYGFYNNSSIIYYSSDNGGQPTAGGSNWPLRGSKGTYWEGGIRAVGVHSPLLKNKGTVCK
ELVHITDWYPTLISLAEGQIDEDIQLDGYDIWETISEGLRSPRVDILHNIDPYTPRQKMAPG
QQAMGSGTLQSSQPSECSTGNCLQEILATATGSPLSLSATWDRTGGTMNGSPCLAKVYGFS
TSQPTHMRGWYLTGIQES

Important Features:

Signal Peptide:

amino acids 1-37

Sulfatases signature 1.

amino acids 120-132

Sulfatases signature 2.

amino acids 168-177

Tyrosine kinase phosphorylation site.

amino acids 163-169

N-glycosylation sites.

amino acids 157-160, 306-309 and 318-321

FIGURE 44

CGGACGCGTGGGTGCGAGTGGAGCGGAGGACCGGAGCGGCTGAGGAGAGAGGAGGCGCGGGC
TTAGCTGCTACGGGGTCCGGCCGCGCCCTCCCGAGGGGGGCTCAGGAGGAGGAAGGAGGAC
CCGTGCGGAGAATGCCTCTGCGCCTGGAGCCTTGCGCTCCCGCTGCTGCTCTCTGGGTGGCAG
GTGGTTTTCGGGAACCGCGGCCAGTGCAAGGCATCACGGGTGTGTTAGCATCGGCACGTCAAGCCT
GGGGTCTGTCTACTATGGAATAAACTGGCCTGCTGCTACGGCTGGAGAAGAAACAGCAAGGG
AGTCTGTGAAGCTACATGCGAACCTGGATGTAAGTTTGGTGAGTGCGTGGGACCAAAACAAAT
GCAGATGCTTTCCAGGATACACCGGGAACCTGCAGTCAAGATGTGAATGAGTGTGGAATG
AAACCCCGGCCATGCCAACACAGATGTGTGAATACACCGGAAGCTACAAGTGCTTTTGCCCT
CAGTGGCCACATGCTCATGCCAGATGCTACGTGTGTGAACCTTAGGACATGTGCCATGATAA
ACTGTCAGTACAGCTGTGAAGACACAGAAGAAGGGCCAGTGCCCTGTGTCCATCCTCAGGA
CTCCGCCCTGGCCCCAAATGGAAGAGACTGTCTAGATATTGATGAATGTGCCCTCTGGTAAAGT
CATCTGTCCCTACAATCGAAGATGTGTGAACACATTTGGAAGCTACTACTGCAAAATGTGACA
TTGGTTTTCGAAGTGAATATATCAGTGGACGATATGACTGTATAGATATAAATGAATGTACT
ATGGATAGCCATACGTGTCAGCCACCATGCCAATTGCTTCAATACCCAGGGTCTTCAAGTG
TAAATGCAAGCAGGGATATAAAGGCAATGGACTTCGGTGTCTGCTATCCCTGAAAATTCTG
TGAAGGAAGTCTCTCAGAGCACCTTGGTACCATCAAAGACAGAATCAAGAAGTTGCTTGCTCAC
AAAAACAGCATGAAAAAGAAGGCCAAAAATTAATAATGTTTACCCAGAACCCACCAGGACTCC
TACCCCTAAGGTGAACTTGCAGCCCTTCAACTATGAAGAGATAGTTTCCAGAGGCGGGAACT
CTCATGGAGGTAAAAAAGGGAATGAAGAGAAATGAAGAGGGGCTTGAGGATGAGAAAAGAG
AAGAGAAAGCCCTGAAGAATGACATAGAGGAGCGAAGCCTGCGAGGAGATGTGTTTTTCCTT
AAGGTGAATGAAGCAGGTGAATTCGGCCTGATTCTGGTCCAAGGAAAGCGCTAACTTCCAA
ACTGGAACATAAAGATTTAAATATCTCGGTTGACTGCAGCTTCAATCATGGGATCTGTGACT
GGAAACAGGATAGAGAAGATGATTTTGACTGGAATCCTGCTGATCGAGATAATGCTATTGGC
TTCTATATGGCAGTTCCGGCCTTGGCAGGTCACAAGAAAGACATTTGGCCGATTGAACTTCT
CCTACCTGACCTGCAACCCCAAAGCAACTTCTGTTTGCTCTTTGATTACCGGCTGGCCGGAG
ACAAAGTCGGGAAACTTCGAGTGTGTTGTGAAAAACAGTAACAATGCCCTGGCATGGGAGAAG
ACCACGAGTGAGGATGAAAAGTGAAGACAGGGGAAAATTGAGTTGTATCAAGGAACTGATGC
TACCAAAAGCATCATTTTGAAGCAGAACGTGGCAAGGGCAAAACCGGCGAAATCGCAGTGG
ATGGCGTCTTGCTGTTTTCAGGCCTTATGTCCAGATAGCCTTTTATCTGTGGATGACTGAATG
TTACTATCTTTATATTGACTTTGTATGTGAGTTCCTGGTTTTTTTGATATTGCATCATAG
GACCTCTGGCATTTTGAATTAAGTACTAGCTGAAAAATTGTAATGTACCAACAGAAATATTATTG
TAAGATGCTTTCTGTATAGATATGCCAATATTTGCTTTAAATATCATATCAGTGTATCT
TCTCAGTCACTTTCTGAATCTTTCNCATTTATATTATAAATNTGGAAANGTCAGTTTATCTC
CCCTCCTCNGTATATCTGATTTGTATANGTANGTGTGATNGCTTCTCTACAAACATTTCTA
GAAATAGAAAAAAGCACAGAGAAATGTTTAACTGTTTGACTCTTATGATACTCTTGGA
AAGTATGACATCAAGAGATAGACTTTTGCTAAGTGGCTTAGCTGGGTCTTTCATAGCCAAAC
TTGTATATTTAATTCTTTGTAATAATAA

1001111.102001

FIGURE 45

MPLPWSLALPLLLSWVAGGFNAASARHHGLLASARQPGVCHYGTKLACCYGWRRNSKGVCE
ATCEPFGCKFGECVGPKNKRCFPGYTGKTCSDVNECGMKPRPCQHRVCVNTHGSKYKCFCLSGH
MLMPDATCVNSRTCAMINCQYSCEDTEEGPQCLCPSSGLRLAPNGRDCLDIDECASGKVICP
YNRRCVNTFGSYYCKCHIGFELQYISGRYDCIDINECTMDSHTCSHHANCFNTQGSFKCKCK
QGYKGNGLRCSAIPENSVKEVLRAPGTIKDRIKKLLAHKNSMKKKAKIKNVTPEPTRTPK
VNLQPFNYEEIVSRGGNSHGGKKGNEEK

Signal peptide:

amino acids 1-21

EGF-like domain cysteine pattern signature.

amino acids 80-91

Calcium-binding EGF-like domains

amino acids 103-124, 230-251 and 185-206

FIGURE 46

GGGAGCTGCTGCTGTGGCTGCTGGTGTGTGCGCGCTGCTCCTGCTCTTGGTGCAGCTGCTG
CGCTTCCTGAGGGCTGACGGCGACCTGACGCTACTATGGCCGAGTGGCAGGGACGACGCC
AGAATGGGAGCTGACTGATATGGTGGTGTGGGTGACTGGAGCCTCGAGTGAATTGGTGAGG
AGCTGGCTTACCAGTTGTCTAAACTAGGAGTTTCTCTTGTGCTGTGAGCCAGAAGAGTGCAT
GAGCTGGAAGGGTGAAAAGAAGATGCCCTAGAGAATGGCAATTTAAAAAGAAAAAGATATACT
TGTTTTTGCCCCCTTGACCTGACCGACACTGGTTCCCATGAAGCGGCTACCAAAGCTGTTCTCC
AGGAGTTTGGTAGAATCGACATTCTGGTCAACAATGGTGGAATGTCCCAGCGTTCTCTGTGC
ATGGATACCAGCTTGGATGTCTACAGAAAGCTAATAGAGCTTAACACTTAGGGACGGTGTC
CTTGACAAAATGTGTTCTGCCTCACATGATCGAGAGGAAGCAAGGAAAGATTGTTACTGTGA
ATAGCATCCTGGGTATCATATCTGTACCTCTTTCATTGGATACTGTGCTAGCAAGCATGCT
CTCCGGGGTTTTTTTAAATGGCCTTCGAACAGAACTTGCCACATACCCAGGTATAATAGTTTC
TAACATTTGCCCAGGACCTGTGCAATCAAATATGTGGAGAATTCCTAGCTGGAGAAGTCA
CAAAGACTATAGCAATAATGGAGACCAGTCCCACAAGATGACAACCAAGTCGTTGTGTGCGG
CTGATGTTAATCAGCATGGCCAATGATTTGAAAGAAGTTTGGATCTCAGAACAACCTTCTTT
GTTAGTAACATATTTGTGGCAATACATGCCAACCTGGGCTTGGTGGATAACCAACAAGATGG
GGAAGAAAAGGATTGAGAACTTTAAGAGTGGTGTGGATGCAGACTCTTCTTATTTTAAAAATC
TTTAAGACAAAACATGACTGAAAAGAGCACCTGTACTTTTCAAGCCACTGGAGGGGAGAAATG
GAAAACATGAAAACAGCAATCTTCTTATGCTTCTGAATAATCAAAGACTAATTTGTGATTTT
ACTTTTTAATAGATATGACTTTTGCTTCCAACATGGAATGAAATAAAAAATAAATAATAAAG
ATTGCCATGAATCTTGCAAAA

1001191-10401

FIGURE 47

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA36343
><subunit 1 of 1, 289 aa, 1 stop
><MW: 32268, pI: 9.21, NX(S/T): 0
MVVVVTGASSGIGEEELAYQLSKLGVSLVLSARRVHELERVKRRCLEENGLKEKDILVLPDL
TDTGSHEAATKAVLQEFGRIDILVNNGGMSQRS LCMDTSLDVYRKLI E LNYLGTVSLTKCVL
PHMIERKQGKIVTVNSILGIISVPLSIGYCASKHALRGFFNGLRTELATYPGIIIVSNICPGP
VQSNIVENSLAGEVTKTIGNNGDQSHKMTTSRCVRLMLISMANDLKEVWISEQPFLLVTYLW
QYMPTWAWWITNKMGGKRIENFKSGVDADSSYFKIFKTKHD
```

Important Features:

Signal Peptide:

amino acids 1-31

Transmembrane domain:

amino acids 136-157

Tyrosine kinase phosphorylation site.

106-113 and 107-114

Homologous region to Short-chain alcohol dehydrogenase

amino acids 80-90, 131-168, 1-13 and 176-185

FIGURE 48

CGGACGTGGGCACCGCCATCAGCTGTTTCGCGCGTCTTCTCCTCCAGGTGGGGCAGGGTTTC
GGGCTGTTGGAGCATGTGCTGGGACAGGACAGCATCCTCAATCAATCCAACAGCATATTCCG
TTGCATCTTCTACACACTACAGCTATTGTTAGGTTGCCTGCGGACACGCTGGGCCCTCTGTCC
TGATGCTGCTGAGCTCCTCGGTGTCTCTCGCTGGTTCTGTCTACCTGGCCTGGATCCTGTTT
TTCGTGCTCTATGATTTCTGCATTGTTTGTATCACACCTATGCTATCAACGTGAGCCTGAT
GTGGCTCAGTTTTCCGGAAGGTCCAAGAACCCAGGGCAAGGCTAAGAGGCACTGAGCCCTCA
ACCCAAGCCAGGCTGACCTCATCTGCTTTGCTTTGGTCTTCAAGCCGCTCAGCGTGCCGTGTG
GACAGCGTGGCCCCGGCCCCCAAGCCTCAGGAGGGCAACACAGTCCTTGGCGAGTGGCCCC
TGGCAGGCCAGTGTGAGGAGGCAAGGAGCCACATCTGCAGCGGCTCCTCGTGGCAGACAC
CTGGGTCTCTCACTGCTGCCCACTGCTTTGAAAAGGCAGCAGCAACAGAACTGAATTCCCTGGT
CAGTGGTCTCTGGGTTCTCTGCAGCGTGAGGGACTCAGCCCTGGGGCCGAAGAGGTGGGGGTG
GCTCCCTGCAGTTGCCAGGGCCTATAACCACTACAGCCAGGGCTCAGACCTGGCCCTGTCT
GCAGCTCGCCCCACCCACGACCCACACACCCCTCTGCCTGCCCCAGCCCGCCCATCGCTTCC
CCTTTGGAGCCTCCTGCTGGGCCACTGGCTGGGATCAGGACACCAAGTGATGCTCCTGGGACC
CTAGCAAATCTGCGCTGCGTCTCATCAGTCGCCCCACATGTAACGTATCTACAACCAAGCT
GCACCAGCGACACCTGTCCAACCCGGCCCCGGCTGGGATGCTATGTGGGGGGCCCCAGCCTG
GGGTGCAGGGCCCCCTGTCAAGGAGATTCCGGGGGCCCTGTGCTGTGCCTCGAGCCTGACGGA
CACTGGGTTCAAGCTGGCATCATCAGCTTTGCATCAAGCTGTGCCCAGGAGGACGCTCCTGT
GCTGCTGACCAACACAGCTGCTCACAGTTCTCTGGCTGCAGGCTCGAGTTTCAGGGGGCAGCTT
TCCTGGCCCCAGAGCCAGAGACCCCGGAGATGAGTGATGAGGACAGCTGTGTAGCCTGTGGA
TCCTTGAGGACAGCAGGTCCCAGGCAGGAGCACCTCTCCCATGCCCCTGGGAGGCCAGGCT
GATGCACCAGGGACAGCTGGCCTGTGGCGGAGCCCTGGTGTGAGAGGAGGGCGGTGCTAACTG
CTGCCCACTGCTTCATTGGGCGCCAGGCCCCAGAGGAATGGAGCGTAGGGCTGGGGACCA
CCGGAGGAGTGGGGCCTGAAGCAGCTCATCTGCATGGAGCCTACACCCACCTGAGGGGGG
CTACGACATGGCCCTCCTGCTGCTGGCCAGCCTGTGACACTGGGAGCCAGCCTGCGGCC
TCTGCCTGCCCTATCCTGACCACCACTGCCTGATGEGGAGCGTGGCTGGGTTCTGGGACGG
GCCCGCCAGGAGCAGGCATCAGCTCCTCCAGACAGTGCCTGACCCCTCTGGGGCCTAG
GGCTGCAGCCGGCTGCATGCAGCTCCTGGGGGTGATGGCAGCCCTATTCTGCCGGGGATGG
TGTGTACCACTGCTGTGGGTGAGCTGCCCAGCTGTGAGGGCCTGTCTGGGGCAACCACTGGTG
CATGAGGTGAGGGGCACATGGTTCTTGGCCGGGCTGCACAGCTTCGGAGATGCTTGCCAAAG
CCCCCGCAGGCCGGCGGTCTTCAACCGCGCTCCTGCCTATGAGGACTGGGTGAGCAGTTTGG
ACTGGCAGGTCTACTTCGCGGAGGAACAGAGCCCGAGGCTGAGCCTGGAAGCTGCCTGGCC
AACAATAAGCCAACCAACAGCTGCTGGACAGGGGACCTGGCCATTCTCAGGACAAGAGAATGC
AGGCAGGCAAATGGCATTACTGCCCTGTCTCTCCCAACCTGTCTGTTGTGATTCCAGGCAC
CAGGGCAGGCCCAGAAGCCAGCAGCTGTGGGAAGGAACCTGCCTGGGGCCACAGGTGCCCA
CTCCCCACCTGCAGACAGAGGGGTGTCTGTGGACACTCCCAACCCAACTCTGCTACCAAGC
AGGCGTCTCAGCTTTCTCTCTCTTTACTCTTTAGATACAATCACGCCAGCCAGCTGTGTTT
TGAAAATTTCTTTTTTGGGGGGCAGCAGTTTTCTTTTTTAAACTTAAATAAATTGTTAC
AAATAAAAA

FIGURE 49

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA40571

MLLSSLVSLAGSVYLAWILFFVLYDFCIVCITYAINVSLMWLSFRKVQEPQSKAKRHGNTV
PGEWFWQASVRRQGAHICSGSLVADTWVLTAAHCFEKAATELNSWSVVLGSLQREGLSPGA
EEVGVAALQLPRAYNHYSQGSDDLALLQLAHPTTHTPLCLPQPAHRFPFGASCWATGWDQDTS
DAPGTLRLNLRLLISRPTCNCIYNQLHQRHLSNPARPGMLCGGPQPGVQGPCQGDSSGGPVLC
LEPDGHWVQAGIISFASSCAQEDAPVLLTNTAAHSSWLQARVQGAAFLAQSPETPEMSDEDS
CVACGSLRTAGPQAGAPSPWPWEARLMHQGQLACGGALVSEEAVLTAAHCFIGRQAPEEWSV
GLGTRPEEWGLKQLILHGAYTHPEGGYDMALLLAQPVTLGASLRPLCLPYPDHHLDPDGERG
WVLGRARPGAGISSLTQVTPVTLTGPRACSRLLHAAPGGDGSFILPGMVCTSAVGELPSCEGLS
GAPLVHEVRGTWFLAGLHSFGDACQGPAPPAVFTALPAYEDWVSSLDWQVYFAEEPEPEAP
GSCLANISQPTSC

Important features:

Signal peptide:

amino acids 1-15

Homologous region to Serine proteases, trypsin family

amino acids 79-95, 343-359 and 237-247

N-glycosylation sites.

amino acids 37-40 and 564-567

Kringle domains

amino acids 79-96, 343-360 and 235-247

FIGURE 50

CGGGCCGCCCCCGCCCCCAATTCGGGCCCGGGCCTCGCTGCGGCGGCGACTGAGCCAGGCTGG
GCCGCGTCCCTGAGTCCCAGAGTCGGCGCGGCGCGGCAGGGGCAGCCTTCCACCACGGGGAG
CCCAGCTGTCAGCCGCCTCACAGGAAGATGCTGCGTCGGCGGGGCAGCCCTGGCATGGGTGT
GCACTGTGGGTGACGCCCTGGGAGCACTGTGGTTCTGCCTCACAGGAGCCCTGGAGGTCCAGG
TCCCTGAAGACCCAGTGGTGGCACTGGTGGGCACCGATGCCACCCTGTGCTGCTCCTTCTCC
CCTGAGCCTGGCTTACGCTGGCACAGCTCAACCTCATCTGGCAGCTGACAGATACCAAACA
GCTGGTGACAGCTTTGCTGAGGGCCAGGACCAGGGCAGCGCCTATGCCAACCGCACGGCCC
TCTTCCCGGACCTGTGGCACAGGGCAACGCATCCCTGAGGCTGACGCGCTGCGTGTGGCG
GACGAGGGCAGCTTACCTGCTTCGTGAGCATCCGGGATTTTCGGCAGCGCTGCCCTCAGCCT
GCAGGTGGCCGCTCCCTACTCGAAGCCAGCATGACCTGGAGCCCAACAAGGACCTGCGGC
CAGGGGACACGGTGACCATCACGTGCTCCAGCTACCAGGGCTACCTGAGGCTGAGGTGTTT
TGGCAGGATGGGCAGGGTGTGCCCCTGACTGGCAACGTGACCAGTCGCAGATGGCCAACGA
GCAGGGCTTGTTTGTATGTGCACAGCGTCTGCGGGTGGTGTGGGTGCGAATGGCACCTACA
GCTGCCTGGTGCGAACCCCGTGTGTCAGCAGGATGCGCACRGCTCTGTCAACCATCACAGGG
CAGCCTATGACATTTCCCCCAGAGGCCCTGTGGGTGACCGTGGGGCTGTCTGTCTGTCTCAT
TGCACTGCTGGTGGCCCTGGCTTTCGTGTGCTGGAGAAAGATCAAAACAGAGCTGTGAGGAGG
AGAAATGCAGGAGCTGAGGACCAGGATGGGGAGGGAGAAGGCTCCAAGACAGCCCTGCAGCCT
CTGAAACACTCTGACAGCAAAGAAGATGATGGACAAGAAATAGCCTGACCATGAGGACCAGG
GAGCTGCTACCCCTCCCTACAGCTCCTACCTCTGGCTGCAATGGGGCTGCACTGTGAGCCC
TGCCCCCAACAGATGCATCCTGCTGACAGGTGGGCTCCTTCTCCAAGGATGCGATACAC
AGACCAGTGTGCAGCCTTATTTCTCAATGGACATGATTTCCCAAGTCATCCTGCTGCCTTTT
TTCTTATAGACACAATGAACAGACCACCCACAACCTTAGTTCTCTAAGTCATCCTGCCTGCT
GCCTTATTTACAGTACATACATTTCTTAGGGACACAGTACACTGACCACATCACCACCTC
TTCTTCCAGTGTGCGTGGACCATCTGGCTGCCCTTTTTTCTCCAAAAGATGCAATATTCAGA
CTGACTGACCCCTGCCTTATTTACCAAAGACACGATGCATAGTACCCCCGGCCTGTGTTT
TCCAATGGCCGTGATACACTAGTGATCATGTTTACGCCCTGCTTCCACCTGCATAGAATCTTT
TCTTCTCAGACAGGGACAGTGCGCCTCAACATCTCCTGGAGTCTAGAAGCTGTTTCTCTTC
CCCTCCTTCTCCTGCCCAAGTGAAGACAGGGGCAGGGCCAGGAATGCTTTGGGGACACCG
AGGGGACTGCCCCCCACCCCCACCATGGTGCTATTCTGGGGCTGGGGCAGTCTTTTCTGGC
TTGCCTCTGGCCAGCTCCTGGCCTCTGGTAGAGTGAGACTTTCAGACGTTCTGATGCCCTCCG
GATGTCTATCTCTCCCTGCCCCAGGAATGGAAGATGTGAGGACTTCTAATTTAAATGTGGGAC
TCGGAGGGATTTTGTAACTGGGGGTATATTTTGGGGAAATAAATGTCTTTGTAAAAAAA
AAAAAAAAAAAAA

FIGURE 51

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA41386

><subunit 1 of 1, 316 aa, 1 stop, 1 unknown

><MW: -1, pI: 4.62, NX(S/T): 4

MLRRRGSPGMGVHVGAAALGALWFCLTGALEVQVPEDPVVALVGTDTATLCCSFSPPEGFSLAQ
LNLIIQLTDTKQLVHSFAEGQDQGSAYANRTALFPDLLAQGNASLRLQVRVVADEGSFTCFV
SIRDFGSAAVSLQVAAPYSKPSMTLEPNKDLRPGDTVTITCSSYQGYPEAEVFWQDGGQGVPL
TGNVTTSQMANEQGLFDVHSVLRVVLGANGTYSCLVRNPVLQQDAHXSVTITGQPMTFPPEA
LWVTVGLSVCLIALLLVALAFVCWRKIKQSCEEENAGAEDQDGEGEKSKTALQPLKHSDSKED
DGQEIA

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 251-270

N-glycosylation site.

amino acids 91-94, 104-107, 189-192 and 215-218

Homologous region to Immunoglobulins and MHC

amino acids 217-234

FIGURE 52

TTCTGACCCCTTGAGAAAAGAGTTGGTGGTAAATGTGCCACGTCCTTCTAAGAAGGGGGAGTC
CTGAACCTTGTCTGAAGCCCTTGTCCGTAAGCCTTGAACCTAGCTTCTTAAATCTATGAAGTCG
AGGGACCTTTCTGCTGCTTTTGTAGGGACTTCTTTCCTTGCTTCAGCAACATGAGGCTTTTCT
TGTGGAACCGGGTCTTGACTCTGTTCGTCACCTTCTTTGATTGGGGCTTTGATCCCTGAACCA
GAAGTGAATTTGAAGTTCTCCAGAAGCCATTTCATCTGCCATCGCAAGACCAAGAGGGGGA
TTTGATGTTGGTCCACTATGAAGGCTACTTAGAAAAGGACGGCTCCTTATTTCACTCCACTC
ACAAACATAACCAATGGTCAGCCCATTTGGTTTACCTGGGCATCTTGAGGGCTCTCAAAGGT
TGGGACCAAGGGCTTGAAGGAATGTGTAGGAGAGAAGAGAAAGCTCATCATTCTCCTCTC
TCTGGGCTATGGAAGAAGGAAAAGGTAAATTTCCCCAGAAAGTACACTGATATTTAATA
TTGATCTCCTGGAGATTGAAATGGACCAAGATCCCATGAATCATTTCCAAGAAATGGATCTT
AATGATGACTGGAACTCTCTAAAGATGAGGTTAAAGCATATTTAAAGAAGGAGTTTGA
ACATGGTGGGTGGTGAATGAAAGTCATCATGATGCTTTGGTGGAGGATATTTTGTATAAG
AAGATGAAGACAAAGATGGGTTTATATCTGCCAGAGAATTTACATATAAACACGATGAGTTA
TAGAGATACATCTACCTTTTAAATATAGCACTCATCTTTCAAGAGAGGGCAGTCATCTTTAA
AGAACATTTTATTTTATACAATGTCTTCTTGCTTTGTTTTTATTTTATATATATTTTT
CTGACTCTCTATTTAAAGAACCCCTTAGGTTTCTAAGTACCCATTCTCTTCTGATAAGTTATT
GGGAAGAAAAAGCTAATTTGGTCTTTGAAATAGAAGACTTCTGGACAATTTTCACTTTACAG
ATATGAAGCTTTGTTTTTACTTTCTCACTTATAAATTTAAATGTTGCAACTGGGAATATACC
ACGACATGAGACCAGGTTATAGCACAATTAGCACCTATATTTCTGCTTCCCTCTATTTTC
TCCAAAGTTAGAGGTCACATTTGAAAAGCCTTTTGCAATAGCCCAAGGCTTGCATTTTCAT
GTTATAATGAAATAGTTTATGTGTAACCTGGCTCTGAGTCTCTGCTTGAGGACAGAGGAAAA
TGGTTGTTGGACCTGACTTGTTAATGGCTACTGCTTTACTAAGGAGATGTGCAATGCTGAAG
TTAGAAACAAGGTTAATAGCCAGGCATGGTGGCTCATGCCCTGTAATCCAGCACTTTGGGAG
GCTGAGGCGGGCGGATCACCTGAGGTTGGGAGTTGAGACCCAGCCTGACCAACACGGAGAAA
CCCTATCTCTACTAAAATACAAAGTAGCCCGCGTGGTGATGCGTGCTGTAATCCCAGCT
ACCCAGGAAGGCTGAGGCGGCAGAATCACTTGAAACCCGAGGCCAGGTTGCGGTAAAGCCGAG
ATCACCTNCAGCCTGGACACTCTGTCTCGAAAAAAGAAAAGAACACGGTTAATACCATATNA
ATATGTATGCATTGAGACATGCTACCTAGGACTTAAGCTGATGAAGCTTGGCTCTAGTGAT
TGGTGGCCTATTATGATAAAATAGGACAAATCATTTATGTGTGAGTTTCTTTGTAATAAAATG
TATCAATATGTTATAGATGAGGTAGAAAGTTATATTTATATTCAATATTTACTTCTTAAGGC
TAGCGGAATATCCTTCTGGTTCTTTAATGGGTAGTCTATAGTATATTATACTACAATAACA
TTGTATCATAAAGATAAAGTAGTAAACCAGTCTACATTTTCCCATTTCTGTCTCATCAAAAC
TGAAGTTAGCTGGGTGGTGGCTCATGCCGTGAATCCCAGCACTTTGGGGGCCAAGGAGG
TGGATCACTTGAGATCAGGAGTTCAAGACCAGCCTGGCAACATGGTGAAACCTTGTCTCTA
CTAAAAATACAAAAATTTAGCCAGGCGTGGTGGTGACACCTGTAGTCCAGCTACTCGGGAG
GCTGAGACAGGAGATTGGCTTGAACCCGGGAGGCGGAGGTTGCGAGTGAGCCCAAGATTGTGCC
ACTGCATCTCAGCCTGGGTGACAGAGCAAGACTCCATCTCAAAAAAAAAAAAAAGAGCAGA
CCTACAGCAGCTACTATTGAATAAATACCTATCCTGGATTTT

```
><subunit 1 of 1, 211 aa, 1 stop
```

MRLFLWNAVLTLFVTSLIGALIPPEPEVKIEVLQKPFICHRKTKGGDMLLVHVEGYLEKDGSL
FHSTHKHNNQPIWFTLGLILEALKGWDQGLKGMCVGEKRKLIIPPALGYGKEGKGKIPPEST
LIFNIDLLEIRNGPRSHESFQEMDLNDDWKLKDEVKAYLKKEFEKHGAVVNESHHDALVED
IFDKEDDKDGFISAREFTYKHDEL

Signal peptide:

amino acids 1-20

amino acids 176-179

amino acids 143-146, 156-159, 178-181 and 200-203

amino acids 208-211

amino acids 78-114 and 118-131

amino acids 191-203, 184-203 and 140-159

amino acids 183-203

FIGURE 54

AATAAAGCTTCCTTAATGTTGTATATGTCCTTTGAAGTACATCCGTGCATTTTTTTTTTAGCAT
 CCAACCATTCCTCCCTTGTAGTTCTCGCCCCCTCAAATCACCCCTCTCCCGTAGCCCAACCGA
 CTAACATCTCAGTCTCTGAAAATGTCACAGAGATGCTGGCTACCTCGCCCTGCCCTCAGCCT
 CACGGGGCTCAGTCTCTTTTTTCTCTTTGGTGCCACGAGCAGGAGCATGGAGGTCAAGTAC
 CTGCCACCCCTCAACGTCTCAATGGCTCTGACGCCCGCCTGCCCTGCACCTTCAACTCCTGC
 TACACAGTGAACCAACAACAGTTCTCCTGAACTGGACTTACAGGAGTGCAACAACCTGCTC
 TGAGGAGATGTTCTCCAGTTCCGCATGAAGATCATTAACCTGAAGCTGGAGCGGTTTCAAG
 ACCGCGTGGAGTTCTCAGGGAACCCAGCAAGTACGATGTGTGGTGTGCTGAGAAACGTG
 CAGCCGGAGGATGAGGGGATTTACAACGTCTACATCATGAACCCCCCTGACCCGCCACCGTGG
 CCATGGCAAGATCCATCTGCAGGTCCTCATGGAAGAGCCCCCTGAGCGGACTCCACGGTGG
 CCGTGATTGTGGGTGCCCTCCGTCGGGGGCTTCTGGCTGTGGTTCATCTTGGTGTGATGGTG
 GTCAAGTGTGTGAGGAGAAAAAAGAGCAGAAGCTGAGCACAGATGACCTGAAGACCGAGGA
 GGAGGGCAAGACGGACGGTGAAGCAACCCGGATGATGGCGCCAAGTAGTGGGTGGCCGGCC
 CTGCAGCTCTCCGTGTCCGCTCTCCTCCCTCTCCGCCCTGTACAGTGACCTGCTGCTCG
 CTCTTGGTGTGCTTCCCGTGACCTAGGACCCAGGGCCACCTGGGGCTCTCTGAACCCCG
 ACTTCGTATCTCCCACCTGCACCAAGAGTGACCCACTCTCTTCCATCCGAGAAACCTGCCA
 TGCTCTGGGACGTGTGGGCCCTGGGGAGAGGAGAGAAAGGGCTCCCACCTGCCAGTCCCTGG
 GGGAGGCAGGAGGCACATGTAGGGTCCCAGAGAGAAGGGAGTGGTGGGCAGGGGTAGA
 GGAGGGGCGCTGTACCTGCCAGTGCTTGCTGGCAGTGGCTTCAGAGAGGACCTGGTGG
 GGAGGGAGGGCTTTCCTGTGCTGACAGCGCTCCCTCAGGAGGGCCTTGGCTGGCACGGCTG
 TGCTCCTCCCTGCTCCAGCCCAGAGCAGCATCAGGCTGGAGGTGACGATGAGTTCTGA
 AACTTGGAGGGGCATGTTAAAGGATGACTGTGCATTCCAGGGCACTGACGGAAAGCCAGGG
 CTGCAGGCAAGCTGGACATGTGCCCTGGCCCAGGAGGCCATGTTGGGCCCTCGTTTCCATT
 GCTAGTGGCCTCCTTGGGGCTCTGTTGGCTCCTAATCCCTTAGGACTGTGGATGAGGCCAG
 ACTGGAAGAGCAGCTCCAGGTAGGGGGCCATGTTTCCCAGCGGGGACCCACCAACAGAGGCC
 AGTTTCAAAGTCAGCTGAGGGGCTGAGGGTGGGGCTCCATGGTGAATGACAGGTTGCTGCAG
 GCTCTGCTCTCTCCAAACAACCTCCCTTCTGGGGACAAAAGTGACAATTGTAGGCCAGGC
 ACAGTGGCTCAGCCTGTAAATCCAGCACTTTGGGAGGCCAAGCGGGTGGATTACTCCAT
 CTGTTTGTAGTAAATGGGCAAAACCCATCTCTACTAAAAATACAAGAATTAGCTGGGCGTG
 GTGGCGTGTGCTGTAAATCCAGCTATTTGGGAGGCTGAGGCAGGAGAATCGCTTGAGCCCG
 GGAAGCAGAGGTTGCACTGAAGTGTAGTGTAGTGCCACTGCAATTTCAGCCTGGGTGAC
 ATAGAGAGACTCCATCTCAAAAAA

1001791.02400

FIGURE 55

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45415

<subunit 1 of 1, 215 aa, 1 stop

<MW: 24326, pI: 6.32, NX(S/T): 4

MHRDAWLPRPAFSLTGLSLFFSLVPPGRSMEVTVPATLNVLNGSDARLPCTFNSCYTVNHKQ
FSLNWTYQECNNCSEEMFLQFRMKIINLKLERFQDRVEFSGNPSKYDVSVMRLRNVQPEDEGI
YNCYIMNPDRHRHGHIHLQVLMEEPPERDSTVAVIVGASVGGFLAVVILVLMVVKCVRRK
KEQKLSTDDLKTEEEGKTGEGNPDDGAK

Important features:

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 161-179

Immunoglobulin-like fold:

amino acids 83-127

N-glycosylation sites.

amino acids 42-45, 66-69 and 74-77

1001191.102401

FIGURE 56

GTTGTATATGTCCTGAAGTACATCCGTGCATTTTTTTTAGCATCCAACCATCCTCCCTTGTA
GTTCTCGCCCCCTCAAATCACCTTCTCCCTTAGCCCCACCCNACTAACATCTCAGTCTCTGAA
AATGCACAGAGATGCCTGGCTACCTCGCCCTGCCTTCAGCCTCACGGGGCTCAGTCTCTTTT
TCTCTTTGGTGCCACCAGGACGGAGCATGGAGGTCCACAGTACCTGNCCACCCTCAACGTCC
TCAATGGCTCTGACGCCCCGCTGCCCTGCCCTTCAACTCCTGCTACACAGTGAACCACAAAC
AGTTCTCCCTGAACTGGACTTACCAGGAGTGCAACAACCTGCTCTGAGGAGATGTTCTCCAG
TTCCGCATGAAGATCATTAACCTGAAGCTGGAGCGGTTTCAAGACCGCGTGGAGTTCTCAGG
GAACCCAGCAAGTACGATGTGTGCGGTGATGCTGAGAAACGTGCAGCCGGAGGATGAGGGGA
TTTACAACCTGCTACATCATGAACCCCCC

10017491.102401

Chickadee

TCACGGGGCTCATCTCTTTTTCTCTTTGGTGCCACCAGGACGGAGCATGGAGGTNCACATA
CCTGCCACCTCAACGTCTCAATGGCTTTGACGCCGCGCTGCCCTGACCTTCAACTCNG
CTACACAGTGAAACCAAAACAGTTCTCCGTGAAGTGGATTATACCAGGAGTGCAACAACATGGC
TCTGAGGAGATGTTCTCCAGTTCCTCCGATGGAAGATCATTTAACTGAAAGCTGGAAGCGG
TTTTCAAGAACCGCTGGAGATTTCTCAGGGAACCCAGCAAGTACGATGTGTCGGTGATGC
TGAGAAACGTGCACGCCGAGATGAGGGGATTTACACTGCTACATCATGAACCCCC

FIGURE 59

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44189

><subunit 1 of 1, 412 aa, 1 stop

><MW: 46658, pI: 6.65, NX(S/T): 4

MGLHLRPYRVGLLPDGLLFLLLLLMLLADPALPAGRHPVVLVPGDLGNQLEAKLDKPTVVH
YLCSKKTESYFTIWLNLLELLPVIIDCWIDNIRLVYNKTSRATQFPDGVDRVVPFGKTFSL
EFLDPSKSSVGSYFHTMVESLVGWGYTRGEDVRGAPYDWRRAPNENGPYFLALREMIEMYQ
LYGGPVLVAHSMGNMYTLYFLQRQPQAWKDKYIRAFVSLGAPWGGVAKTLRLVLSGDNNRI
PVIGPLKIREQQRSASVSTSWLLPYNYTWSPEKVVFVQTPTINYTLRDYRKFFQDIGFEDGWLM
RQDTEGLVEATMPPGVQLHCLYGTGVPDPSFYYESFFDRDPKICFGDGDGTNVLKSALQCQ
AWQSRQEHQVLLQELPGSEHIEMLANATTLAYLKRVLG

Important features:

Signal peptide:

amino acids 1-28

Potential lipid substrate binding site:

amino acids 147-164

N-glycosylation sites.

amino acids 99-102, 273-276, 289-292 and 398-401

Lipases, serine proteins

amino acids 189-201

Beta-transducin family Trp-Asp repeat

amino acids 353-365

FIGURE 60

CGGACGCGTGGGCGGACGCGTGGGGCGGCGGAGCGGCGGCGACATGGAGAGCGGG
GCCTACGGCGCGGCCAAGGCGGGCGGCTCTTCGACCTGCGGCGCTTCTGACGCAGCCGCA
GGTGGTGGCGCGCGCGTGTGCTTGGTCTTCGCCTTGATCGTGTTCTCCTGCATCTATGGTG
AGGGCTACAGCAATGCCACGAGTCTAAGCAGATGTACTGCGTGTTCAACCGCAACGAGGAT
GCCTGCCGCTATGGCAGTGCCATCGGGGTGCTGGCCTTCTTGGCCTCGGCCCTTCTTCTGGT
GGTCGACGCGTATTTCCCCCAGATCAGCAACGCCACTGACCGCAAGTACCTGGTCATTGGTG
ACCTGCTCTTCTCAGCTCTCTGGACCTTCTGTGGTTTGTGGTTTCTGCTTCTCACCAC
CAGTGGGCAGTCACCAACCCGAAGGACGTGCTGGTGGGGCCGACTCTGTGAGGGCAGCCAT
CACCTTCAGCTTCTTTTCCATCTTCTCCTGGGGTGTGCTGGCCTCCCTGGCCTACCAGCGCT
ACAAGGCTGGCGTGGACGACTTCATCCAGAATTACGTTGACCCCACTCCGGACCCCAACACT
GCCTACGCCCTCTACCCAGGTGCATCTGTGGACAACCTACCAACAGCCACCCTTACCCAGAA
CGCGGAGACCACCGAGGGCTACCAGCCGCCCTGTGTACTGAGTGGCGGTTAGCGTGGGAA
GGGGACAGAGAGGGCCCTCCCTCTGCCCTGGACTTTCCTATCAGCCTCCTGGAAGTGCCA
GCCCCCTCTCTTTACCTGTTCCATCCTGTGCAGCTGACACACAGCTAAGGAGCCTCATAGCC
TGGCGGGGGCTGGCAGAGCCACACCCCAAGTGCCTGTGCCAGAGGGCTTCAGTCAGCCGCT
CACTCCTCCAGGCACTTTTAGGAAAGGGTTTTAGCTAGTGTTTTCTCGCTTTTAAATGA
CCTCAGCCCCGCTGCAGTGGCTAGAAGCCAGCAGGTGCCCATGTGCTACTGACAAGTGCCT
CAGCTTCCCCCGGCCCGGTGAGGCGGTGAGGAGCCGCTATTATCTGCGTTCTCTGCCAAAG
ACTCGTGGGGCCATCACACCTGCCCTGTGCAGCGAGCCGACAGGCTCTTGTGTCCTCA
CTCAGTTTGTCTTCCCTGTGCCCACTGCTGTATGATCTGGGGGCCACCACCTGTGCCGGT
GGCCTCTGGGCTGCCTCCCGTGGTGTGAGGGCGGGGCTGGTGCTATGGCACTTCTCCTTG
CTCCACCCCTGGCAGCAGGAAGGGCTTTCCTGACAACCCAGCTTTATGTAATATTC
TGCAGTTGTTACTTAGGAAGCCTGGGGAGGGCAGGGGTGCCCATGGCTCCCAGACTCTGTC
TGTGCCGAGTGATTATAAAATCGTGGGGAGATGCCCGGCTGGGATGCTGTTGGAGACG
GAATAAATGTTTTCTCATTCAAAG

FIGURE 61

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48304

<subunit 1 of 1, 224 aa, 1 stop

<MW: 24810, pI: 4.75, NX(S/T): 1

MESGAYGAAGAGGSFDLRRFLTQPQVVARAVCLVFALIVFSCIYGEYSNAHESKQMYCVFN
RNEDACRYGSAIGVLAFLASAFFLVVDAYFPQISNATDRKYLVI GDLLFSALWTF LWVFGFC
FLTNQWAVTNPKDVLVGADSVRAAITFSFFSIFSWGLASLAYQRYKAGVDDFIQNYVDPTP
DPNTAYASYPGASVDNYQQPPFTQNAETTEGYQPPPVY

Important features:

Type II Transmembrane domain:

amino acids 1-45

Other transmembrane domains:

amino acids 74-90, 108-126 and 145-161

N-glycosylation site.

amino acids 97-100

FIGURE 62

GAGCCACTACCTGCTCGAGGGCAGGCCCTGCAGGGCTCATCGGCCAGAGGGTGATCAGTGAGCAGAAGCATG
 CCCGTGGCCGAGGCCCTCCAGGTGGCTGGCGGGCAGGGGACGGAGGGTGATGGCGAGGAAGCGGAGCCAGAGGGG
 ATGTTCAAGGCTCTGAGAGCTCCAGAGAAAGCCCGGGCTACTCCGCTGGTGTCCTGTTTGTGCTGCTG
 GCCTGCTCGTGCTGGCTTCGGCGGGGGTGCTACTCTGGTATTTCTTAGGGTACAAGGCCGAGGTGATGGTCAGC
 CAGGTGTACTCAGGCAGTCTGGTGTACTCAATCGCCACTTCTCCAGGATCTTACCCGCGCGGGAATCTAGTGGC
 TTCCGCAGTGAAACCGCCAAAGCCCGAAGATGCTCAAGGAGCTCATCACCAGCACCCGCTGGGAACTTACTAC
 AACTCCAGCTCCGTCTATTTCCTTTGGGGAGGGACCCCTCACCTGCTTCTTCTGGTTCATTCTCCAAATCCCCGAG
 CACCGCCGGCTGATGCTGAGCCCCGAGGTGGTGAGGCACCTGCTGGTGGAGGAGCTGCTGTCCACAGTCAACAGC
 TCGGCTGGCGTCCCTACAGGGCCGAGTACGAAGTGGACCCCGAGGGCTATGGTATCCTGGAAGCCAGTGTGAA
 GACATAGCTGCATTGAATTCACCGCTGGGTGTGTACCGCTACAGCTACGTTGGGCCAGGGCCAGGTCTCCGCGTG
 AAGGGCCCTGACCACCTGCGCTCCAGCTGCTGTTGGCACCTGCAAGGCCCAAGGACCTCATGCTCAAATCCCCG
 CTGGAGTGGACGCTGGCAGAGTGCAGGGACCGACTGGCCATGTATGACGTGGCGGGGCCCTGGAGAAGAGGCTC
 ATCACTCGGTGTACGGCTGCAGCCCGCAGGAGCCGCTGGTGGAGGTTTGGCGCTCGGGGCCCATCATGGCGCTC
 GTCTGGAAGAAGGGCTGCACAGCTACTACGACCCCTTCTGCTGCTCCGTGCAGCCGGTGGTCTTCCAGGCCGT
 GAAGTGAACCTGACGCTGGACAACAGGCTCGACTCCAGGGCGTCTCAGCACCCCGTACTTCCCGAGCTACTAC
 TCGCCCCAAACCACTGCTCCTGGCACTCAGCGTGCCCTCTCTGGACTACGCGCTTGGCCCTCTGGTTTGTATGCC
 TATGCACTGAGGAGGCGAGAAGTATGATTTGCCGTGCACCCAGGGCCAGTGGAGATCAGAAACAGGAGGCTGTGT
 GGCTTGGCATCTCTGCAGCCCTACGCCGAGAGGATCCCGTGGTGGCCAGGCCGGGATCACCATCAACTTCACC
 TCCAGATCTCCCTACCGGGGCCGGTGTGCGGGTGCACTATGGCTGTACAACAGTCCGAGCCCTCGCCCTGGA
 GAGTTCTCTGTTCTGTGAATGGACTCTGTGTCCTGCTGTGATGGGGTCAAGGACTGCCCAACGGCCCTGGAT
 GAGAGAACTCGTTTGCAGAGCCACTTCCAGTGCAAGAGGACAGCACATGCATCTCATGTGCCCAAGGTCTGT
 GATGGGCAGCTGATTGTCTCAACGGCAGCGATGAAGAGCAGTGCCAGGAAGGGGTGCCATGTGGGACATTACCC
 TTCAGTGTGAGGACCGGAGTGCCTGAAGAAGCCCAACCGCAGTGTGATGGGGGCCCGGACTGCAGGAGCGGC
 TCGGATGAGGAGCACGTGACTGTGGCTCCAGGCCCTTCACGCGCATTTGTGGAGTGTGTCTCCGAG
 GGTGAGTGGCCATGGCAGGCAGCTCCAGTTTCGGGGTGCACACATCTGTGGGGGGCCCTCATGCTGCACCCG
 TGGTGATTAACAGCTGCCACTGCTTCCAGAGGACAGCATGGCTCCACCGCTGTGTGGACCGTGTCTTGGGC
 AAGGTGTGGCAGAACTCGCGCTGCGCTCGAGAGGTGTCTTCAAGGTGAGCCGCTGCTCTGCACCCCTACCA
 GAAGAGGACAGCATGACTGACAGCTGGCGCTGTGCAGCTCGAGCACCGCGTGGTGGCTCGGCGCTGCGCGC
 CCGCTCTGCTGGCCGCGCTCCCTCCACTTCTTCAGGCCGGCTGCAGCTGTGATTTACGGCTGTGGCGCTGTG
 CGGAGGGCGGCCCATCAGCAAGCTCTGCGAGAAAGTGGATGTGATCCACAGGACTGTGCAGCGAG
 GCTATCTGCTACAGGTGAGGCCACGATGCTGTGTGCCGGCTACCGCAAGGGCAGAGATGCTGTACGGGT
 GACTCAGGTGGTCCGCTGGTGTGCAAGGCACTCAGTGGCGCTGCTTCTGGCGGGGCTGGTCACTGGGGCCTG
 GCTGTGGCGCGCTTAACACTTCTGGCGTCTACCCCGCATCACAGGTGATCAGCTGGATCCAGCAAGTGGTG
 ACCTGAAGAACTGCCCCCTGCAAGCAGGGCCCACTCTCGACTCAGAGAGCCAGGGCACTGCCAAGCAGG
 GGGACAAGTATTTCTGGCGGGGTGGGGGAGAGAGCAGGCCCTGTGGTGGCAGGAGGTGGCATTTGTCTCGTCC
 CTGATGCTGCTCTCAGTGATGGCAGGAGGATGGAGAAGTGCAGAGCTGCGGGTCAAGACGTCCCTGAGGACC
 CAGGCCACACCCAGCCCTTCTGCTCCCAATTTCTCTCTCTCCGTCCCTTCTCCACTGCTGCTAATGCAAG
 GCAGTGGCTCAGCAGCAAGATGCTGTGTTACATCCCGAGGAGTGTCTGAGTGTGCGCCCACTCTGTACAGAGG
 CTGTTTGGGCGAGCTTGGCTCCAGAGAGCAGATTCAGCTTGGGAAGCCCTGGTCTAACTTGGGATCTGGGAAT
 GGAAGGTGCTCCATCGGAGGGGACCTCAGAGCCCTGGAGACTGCCAGGTGGGCTGTGCCACTGTAAAGCCAA
 AAGGTGGGGAAGTCTGACTCAAGGCTCTTGCCTCCAGCTGCTGCACTGCGCCCTCACAGGCCAGACCCCT
 CATCTGGAGGTGAGCTCAGCTGCCTTTGGAATAAAGCTGCTGATCAAAAAAAAAAAAAAAAAAAAA

10017191-102401

FIGURE 63

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49152
><subunit 1 of 1, 802 aa, 1 stop
><MW: 88846, pI: 6.41, NX(S/T): 7
MPVAEAPQVAGQGDDGGDEAEPEGMFACEDSKRKARGYLRLVPLFVLLALLVLASAGVL
LWYFLGYKAEVMVVSQVYSGSLRVLNRHFSQDLTRRESSAFRSETAKAQKMLKELITSTRLGT
YYNSSSVYSFGEGPLTCFFWFILQIPEHRRMLLSPEVVQALLVEELLSTVNSSAAVPYRAEY
EVDPEGLVILEASVKDIAALNSTLGCYRYSYVGQGVRLKGPDLHASSCLWHLQGPDLML
KLRLEWTLAECDRLAMYDVAGPLEKRLITSVYGCSRQEPVVEVLASGAIMAVVWKKGLHSY
YDPFVLSVQPVVVFQACEVNLTLNRLDSQGVLSPTYFPSPYSPQTHCSWHLTVPSLDYGLAL
WFDAYALRRQKYDLPTCQGQWTIQNRRLCGLRILQPYAERIPVVATAGITINFTSQISLTGP
GVRVHYGLYNQSDPCPGEFCLSVNGLCVPACDGVKDCFNGLDERNCVCRATFQCKEDSTCIS
LPKVCDDGQPDCLNGSDEEQCQEGVPCGTFTFQCEDRSCVKKPNPQCDGRPDCCRDSDEEHCD
CGLQGPSSRIVGGAVSSEGEWPWQASLQVRGRHICGGALIADRVVITAAHCFQEDSMASVTL
WTVFLGKVVQNSRWPGEVSFVKVSRLLHPYHEEDSHDYDVALLQLDHPVVRSAAVRPVCLPA
RSHFFEPGLHCWITGWGALREGGPISNALQKVDVQLIPQDLCEAYRYQVTPRMLCAGYRKG
KKDACQGDGSGGPLVCKALSGRWFLAGLVSWGLGCGRPNYFGVYTRITGVISWIIQQVVT

Important features:

Type II transmembrane domain:

amino acids 46-67

Serine proteases, trypsin family, histidine active site.

amino acids 604-609

N-glycosylation sites.

amino acids 127-130, 175-178, 207-210, 329-332, 424-427, 444-447
and 509-512

Kringle domains.

amino acids 746-758 and 592-609

Homologous region to Kallikrein Light Chain:

amino acids 568-779

Homologous region to Low-density lipoprotein receptor:

amino acids 451-567

FIGURE 64

GCACCCAGGGCCAGTGGACGATCCAGAACAGGAGGCTGTGTGGCTTGCGCATCCTGCAGCCC
TACGCCGAGAGGATCCCCGTGGTGGCCACGGCCGGGATCACCATCAACTTCACCTCCCAGAT
CTCCCTCACCGGGCCGGTGTGCGGGTGCACTATGGCTTGTACAACCAGTCCGAGCCCCTGCC
CTGGAGAGTTCTCTGTTCCTGTGAATGGACTCTGTGTCCCTGCTGTGATGGGGTCAAGGAC
TGCCCCAACGGCCTGGATGAGAGAAACTGCGTTTGAGAGCCACATTCCAGTGCAAGAGGA
CAGCACATGCATCTCACTGCCCAAGGTCTGTGATGGGCAGCCTGATTGTCTCAACGGCAGCG
ATGAAGAGCAGTGCCAGGAAGGGGTGCCATGTGGGACATTACCTTCCAGTGTGAGGACCGG
AGCTGCGTGAAGAAGCCCAACCCGAGTGTGATGGGCGGCCGACTGCAGGGACGGCTCGGA
TGAGGAGCACTGTGACTGTGGCCTCCAGGGCCCCCTCCAGCCGATTGTTGGTGGAGCTGTGT
CCTCCGAGGGTGAGTGGCCATGGCAGGCCAGCCTCCAGGTTCCGGGGTCGACACATCTGTGGG
GGGGCCCTCATCGCTGACCGCTGGGTGATAACAGCTGCCCACTGCTTCCAGGAGGACAGCAT
GGCCTCCACGGTGTCTGTGGACCGTTCCTGGGCAAGGTGTGGCAGAACTCGCGCTGGCCCTG
GAGAGGTGTCTTCAAGGTGAGCGCCCTGCTCCTGCACCCGTACCACGAAGAGGACAGCCAT
GACTACGACGTGGCGCTGTCTGCAGCTCGACCACCCGGTGGTGCCTCGGCCCGCTGCGCCC
CGTCTGCTGCCCGCGCGCTCCCACTTCTTCGAGCCCGCTGCAGTGTGATTACGGGCT
GGGGCGCCTTGCGCGAGGGCGGCCCCATCAGCAACGCTCTGCAGAAAGTGGATGTGCAGTTG
ATCCACAGGACCTGTGCAGCGAGGCCCTATCGCTACCAGGTGACGCCACGCATGCTGTGTGC
CGGCTACCGCAAGGGCAAGAAGGATGCTGTGAGGTGACTCAGGTGGTCCGCTGGTGTGCA
AGGCACTCAGTGGCCGCTGGTTCTTGGCGGGGCTGGTCACTGGGGCTGGGCTGTGGCCGG
CCTAACTACTTCGGCGTCTACACCCGCATCACAGGTGTGATCAGCTGGATCCAGCAAGTGGT
GACCTGAGGAAGTGGCCCCCTGCAAAGCAGGGCCACCTCCTGGACTCAGAGAGCCAGGGC
AACTGCCAAGCAGGGGGACAAGTAT

FIGURE 65

GGACGAGGGCAGATCTCGTTCTGGGGCAAGCCGTTGACACTCGCTCCCTGCCACCGCCCGGG
CTCCGTGCCGCCAAGTTTTTCATTTTCCACCTTCTCTGCCTCCAGTCCCCCAGCCCCGTGGCCG
AGAGAAGGGTCTTACCGGCCGGGATTGCTGGAAACACCAAGAGGTGGTTTTTGTTTTTTAA
ACTTCTGTTTCTTGGGAGGGGTGTGGCGGGGCAGGATGAGCAACTCCGTTCTCTGTCTCTG
TTTCTGGAGCCTCTGCTATTGCTTTGCTGCGGGGAGCCCCGTACCTTTTGGTCCAGAGGGAC
GGCTGGAAGATAAGCTCCACAAACCAAGCTACACAGACTGAGGTCAAACCATCTGTGAGG
TTTAACCTCCGCACCTCCAAGACCCAGAGCATGAAGGATGTACCTCTCCGTGCGGCCACAG
CCAGCCCTTAGAAGACTGCAGTTTCAACATGACAGCTAAAACTTTTTTCATCATTACGGAT
GGACGATGAGCGGTATCTTTGAAAACTGGCTGCACAACTCGTGTGAGCCCTGCACACAAGA
GAGAAAGACGCGCAATGTAGTTGTGGTTGACTGGCTCCCCCTGGCCCACACAGCTTTACACGGA
TGCGGTCAATAATACCAGGGTGGTGGGACACAGCATTGCCAGGATGCTCGACTGGCTGCAGG
AGAAGGACGATTTTTCTCTCGGGAATGTCCACTTGATCGGCTACAGCCTCGGAGCGCACGTG
GCCGGGTATGCAGGCAACTTCGTGAAAGGAACGGTGGGCCGAATACAGGTTTGGATCCTGC
CGGGCCCATGTTTGAAGGGGCGCATCCACAAGAGGCTCTCTCCGGACGATGACAGATTTTG
TGGATGTCTCCACACCTACACGCGTTCTTCCGCTTGAGCATTGGTATTAGATGCCTGTG
GGCCACATTGACATCTACCCCAATGGGGGTGACTTCAGCCAGGCTGTGGACTCAACGATGT
CTTGGGATCAATTGCATATGGAACAATCACAGAGGTGGTAAAAATGTGAGCATGAGCGAGCCG
TCCACCTCTTGTGACTCTCTGGTGAATCAGGACAAGCCGAGTTTGCCTTCCAGTGCAC
GACTCCAATCGCTTCAAAAAGGGGATCTGTCTGAGCTGCCGAAGAACCGTTGTAATAGCAT
TGGCTACAATGCCAAGAAAATGAGGAACAAGAGGAACAGCAAAATGTACCTAAAAACCCGGG
CAGGATGCCTTTCAGAGGTAACTTCAGTCCCTGGAGTGTCCCTGAGGAAGGCCCTTAATA
CCTCCTTCTTAATACCATGCTGCAGAGCAGGGCACATCCTAGCCCAGGAGAAGTGGCAGCA
CAATCCAATCAAATCGTTGCAAATCAGATTACACTGTGCATGTCTAGGAAAGGGAATCTTT
ACAAAAATAACAGTGTGGACCCCTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAA

```
><subunit 1 of 1, 354 aa, 1 stop
```

MSNSVPLLCFWSLCYCAAGSPVPFPGPEGRLEDKLHKPKATQTVEKPSVRFNLRTSKDPEHE
GCYLSVGHSQPLEDCSFNMTAKTFPIIHGWTMSGIFENWLHLKLVSAHLTREKDVANVVVDWL
PLAHQLYTDVANNTRVVGHSIARMLDWLQEKDDFSLGNVHLIGYSLGAHVAGYAGNFVKGTV
GRITGLDPAGPMFEGADIHKRLSPDDADFVDVLHTYTRSFQSLSIGIQMPVGHIDIYPNGGDF
CQCGCLNDVLGSIAYGTITEVVKCEHRAVHLFVDSLNVQDKPSFAFQCTDSNRKFKKICLS
CRKNCNCSIGYNAKKMRKNRKNCSMYLKTFRAGMFPRGNLOSLECP

Signal peptide:

amino acids 1-16

amino acids 163-172

amino acids 80-83 and 136-139

FIGURE 68

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49631

<subunit 1 of 1, 713 aa, 1 stop

<MW: 76193, pI: 5.42, NX(S/T): 4

MLLATLLLLLLGGALAHDPRIIFPNHACEDPPAVLLEVQGTLQRPLVRDSRTSPANCTWLIL
GSKEQTVTIRFQKLHLACGSRRLTLRSPLQLISLCEAPPSPLQLPGGNVTITYSYAGARAP
MQQGFLLSYSQDWLMCLQEEFQCLNHRCVSAVQRCDGVDACGDGSDEAGCSSDPFPGGLTPRP
VPSLPCNVLTLEDFYGVFSSPGYTHLASVSHPQSCHWLLDHPDGRRLAVRFTALDLGFGDAVH
VYDGPPESSRLLRSLTHFSNGKAVTVETLSGQAVVSYHTVAWSNGRGNATYHVRGYCLP
WDRPCGLGSLGAGEGLGERCYSEAQRCDGSWDCADGTDEEDCPGCPPGHFPCGAAGTSGAT
ACYLPADRCNYQTFCADGADERRCRHCQPGNFRCRDEKCVYETWVCDGQPD CADGSDEWDCS
YVLPKRVITAAVIGSLVCGLLLVIALGCTCKLYAIRTQEYSIFAPLSRMEAEIVQQAPP
SYGQLIAQGAIPPVEDFPTENPNDNSVLGNLRSLLQILRQDMTPGGGPGARRRQRGLMRRLVR
RLRRWGLLPTNTPTPARSEARSQVTPSAAPLEALDGGTGPAREGGAVGGQDGEQAPPLPIKA
PLPSASTSPAPTTVPEAPGPLPSLPLEPSLLSGVVQALRGRLLPSLGGPPGPTSPPGPHTA
VLALEDDDDVLLVPLAEFPGVWVAEAEDEPLLT

Important features:

Signal peptide:

amino acids 1-16

Transmembrane domain:

amino acids 442-462

LDL-receptor class A (LDLRA) domain proteins

amino acids 411-431, 152-171, 331-350 and 374-393

CGAGCTGGGCGAGAAAGTAGGGGAGGGCGGTGCTCCGCCGCGGTGGCGGTTGCTATCGCTTCG
CAGAACCTACTCAGGCAGCCAGCTGAGAAGAGTTGAGGGAAAGTGCTGCTGCTGGGTCTGCA
GACGCGATGGATAACGTGCAGCCGAAAATAAAACATCGCCCTTCTGCTTCAGTGTAAGG
CCAGTGGAAGATGCTGCGGCTGGCACTAACTGTGACATCTATGACCTTTTTTATCATCGCAC
AAGCCCCGAACCATAATATTGTTATCACTGGATTGAAGTCACCGTTATCTTATTTTTCATA
CTTTTATATGTACTCAGACTTGATCGATTAATGAAGTGGTTATTTTGGCCCTTGCTTGATAT
TATCAACTCACTGGTAACAACAGTATTATGCTCATCGTATCTGTGTGGCACTGATACCAG
AAACCACAACATTGACAGTTGGTGGAGGGGTGTTTGCACTTGTGACAGCAGTATGCTGTCTT
GCCACGCGGGCCCTTAATTACCGAAGCTCTGTTCAATCCGACGGTCCCTTACCAGAAAAA
CGCTGTGCATGAAAAAAGAGTTTGTTAATTTTATATTACTTTTAGTTTGATCTAAGT
ATTAACCATATTTCGTATTCTTCCAAAAAATAAAAAA

FIGURE 70

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49645

><subunit 1 of 1, 152 aa, 1 stop

><MW: 17170, pI: 9.62, NX(S/T): 1

MDNVQPKIKHRPFCFSVKGHVKMLRLALTVTSMTEFFIIAQAPEPYIVITGFVTVILFFILL
YVLRDLRLMKWLFWPLLDIINSLVTTVFMLIVSVLALIPETTTTLTVGGGVFALVTVACCLAD
GALIYRKLLFNPSGPYQKKPVHEKKEVL

Important features:

Potential type II transmembrane domain:

amino acids 26-42

Other potential transmembrane domain:

amino acids 44-65, 81-101 and 109-129

Leucine zipper pattern

amino acids 78-99 and 85-106

N-myristoylation site.

amino acids 110-115

Ribonucleotide reductase large subunit protein

amino acids 116-127

FIGURE 71

GGGCGAGAAGTAGGGGAGGGCGTGTCCGCCGCGGTGGCGGTGCTATCGTTTTGCAGAACC
TACTCAGGCAGCCAGNTGAGAAGAGTTGAGGGAAAGTGCTGCTGCTGGGTCTGCAGACGCGA
TGGATAACGTGCAGCCGAAAAATAAACATCGCCCCCTTCTGCTTCAGTGTGAAAGGCCACGTG
AAGATGCTGCGGCTGGCACTAACTGNGACATCTATGACCTTTTTTATNATCGCACAAAGCCCC
TGAACCATAATTGTTATCACTGGATTTGAAGTCACCGTTATCTTATTTTTCATACTTTTAT
ATGTACTCAGACTTGATCGATTAATGAAGTGTTATTTTGGCCTTTGCTTGATATTATCAAC
TCACTGGTAACAACAGTATTTCATGCTCATCGTATCTGTGTTGGCACTGATACCAGAAACCAC
AACATTGACAGTTGGTGGAGGGGTGTTTGCACCTTGTGACAGCAGTATGCTGTNNTGCCGAC

10017191.102401

FIGURE 72

CAGCCCCGCGCGCCGGCCGAGTCGCTGAGCCGCGGCTGCCGGA CGGGA CGGGA CCGGCTAGG
CTGGGCGCGCCCCCGGGCCCCGCGTGGGCATGGGCGCACTGGCCCCGGGCGCTGCTGCTGC
CTCTGCTGGCCAGTGGCTCCTGCGCGCCGCCCGGAGCTGGCCCCCGCGCCCTTCAACGCTG
CCCCCTCCGGGTGGCCGCGGCCACGAACCGCGTAGTTGCGCCACCCCGGGACCCGGGACCCC
TGCCGAGCGCCACGCCGACGGCTTGGCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGG
GCGCCGCCAACTTCTTGGCCATGGTAGACAACCTGCAGGGGCACTCTGGCCGCGGCTACTAC
CTGGAGATGCTGATCGGGACCCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAG
TAACTTTGGCGTGGCAGGAACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGT
CTAGCACATACCGCTCCAAGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACG
GGCTTCGTTGGGGAAGACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAA
CATTGCCACTATTTTTGAATCAGAGAATTTCTTTTTGCTGGGATTAATGGAATGGAATAC
TTGGCTTAGCTTATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCC
CTGGTGACACAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCC
CGTTGCTGGATCTGGGACCAACGAGGTAAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGT
ATAAAGGAGACATCTGTTATACCCCTATTAAAGGAAGAGTGGTACTACCAGATAGAAATTTCTG
AAATTGGAAATTGGAGGCCAAAGCCTTAATCTGGAATGCGAGAGTATAACGCGAGACAAGGC
CATCGTGGACAGTGGCACACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTGG
AAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTTCTGATGGTTTCTGGACTGGGTCCCAG
CTGGCGTGTGGACGAATTCGGAACACCTTGGTCTTACTTCCCTAAAAATCTCCATCTACCT
GAGAGACGAGAATCCAGCAGGTCAATTCGTATCACAATCCTGCCCTCAGCTTTACATTGAGC
CCATGATGGGGGCGGGCTGAATTATGAATGTTACCGATTGCGCATTTCCCCATCCACAAAT
GCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGACAGAGCCAGAA
GAGGGTGGGCTTCGACGAGCGCCCTGTGCAGAAATTCAGGTGCTGCAGTGTCTGAAATTT
CCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCCGCTCAGTCTTTGAGCGAG
CCCATTTTGTGGATTGTGTCTATGCGCTCATGAGCGTCTGTGGAGCCATCCTCCTTGTCTT
AATCGTCTGCTGCTGCTGCCGTTCGGTGTGACGCTCGCCCCGTGACCTGAGGTGCTCA
ATGATGAGTCTCTCTGGTCAGACATCGCTGGAATTGAATAGCCAGGCCTGACCTCAAGCAA
CCATGAACCTCAGCTATTAAAGAAATCACATTTCCAGGGCAGCAGCCGGGATCGATGGTGGCG
CTTTCTCCTGTGCCACCCGTCTTCAATCTCTGTTCTGCTCCAGATGCCTTCTAGATTAC
TGTCTTTTGATTCTTGATTTTCAAGCTTTCAAATCCTCCCTACTTCCAAGAAAAATAATTA
AAAAAAAATTCATTCTAA

FIGURE 73

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45493
><subunit 1 of 1, 518 aa, 1 stop
><MW: 56180, pI: 5.08, NX(S/T): 2
MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGLAL
ALEPALASPAGAAANFLAMVDNLQGDSGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAGTPHS
YIDTYFDTERSSTYRSKGFDTVVKYTQGSWTGFGVGEDLVTIPKGFNTSFLVNIATIFESNF
FLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVAGSGTNGGS
LVLGGIEPSLYKGDWYTPIKEEWYYQIEILKLEIGGQSLNLDCREYNADKAIVDSGTTLLR
LPQKVFDVAVEAVARASLIPEFSDGFWTGSQACWTNSETPWSYFPKISIIYLRDENSRSRFR
ITILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFDRAQKRVGFAASPCA
ETAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCGAILLVLLVLLLPFRRC
QRRPRDPEVVNDESSLVRHRWK

Important features:

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 466-494

N-glycosylation sites.

amino acids 170-173 and 366-369

Leucine zipper pattern.

amino acids 10-31 and 197-118

Eukaryotic and viral aspartyl proteases

amino acids 109-118, 252-261 and 298-310

FIGURE 74

CGCCTCCGCCTTCGGAGGCTGACGCGCCCGGGCGCCGTTCCAGGCCTGTGCAGGGCGGATCG
GCAGCCGCCTGGCGGCGATCCAGGGCGGTGCGGGGCTGGCGGGAGCCGGGAGGCGCGGCC
GGCATGGAGGCGCTGCTGCTGGGCGCGGGGTGCTGCTGGGCGCTTACGTGCTTGCTACTA
CAACCTGGTGAAGGCCCGCCGTGCGGCGGCATGGGCAACCTGCGGGGCGCACGGCCGTGG
TCACGGGCGCCAACAGCGGCATCGGAAAGATGACGGCGCTGGAGCTGGCGCGCCGGGGAGCG
CGCGTGGTGCTGGCCTGCCGACGCCAGGAGCGCGGGGAGGCGGCTGCCCTTCGACCTTCGCCA
GGAGAGTGGGAACAATGAGGTCATCTTCATGGCCTTGGACTTGGCCAGTCTGGCCTCGGTGC
GGGCCTTTGCCACTGCCTTTCTGAGCTCTGAGCCACGGTTGGACATCCTCATCCACAATGCC
GGTATCAGTTCCTGTGGCCGACCCGTGAGGCGTTTAACTGCTGCTTCGGGTGAACCATAT
CGGTCCCTTTCTGCTGACACATCTGCTGCTGCCTTGCTGAAGGCATGTGCCCTAGCCGCG
TGGTGGTGGTAGCCTCAGCTGCCCCTGCTCGGGGACGTCTTGACTTCAAACGCTGGACCGC
CCAGTGGTGGGCTGGCGGCAGGAGCTGCGGGCATATGCTGACACTAAGCTGGCTAATGTACT
GTTTGGCCCGGAGCTCGCCAACAGCTTGAGGCCACTGGCGTCACCTGCTATGCAGCCACC
CAGGGCCTGTGAACTCGGAGCTGTTCTCCTGCGCCATGTTCTCGATGGCTGCGCCACTTTTG
CGCCCATTTGGCTTGGCTGGTCTCCGGGACCAAGAGGGGTGCCAGACACCCCTGTATTG
TGCTCTACAAGAGGGCATCGAGCCCTCAGTGGGAGATATTTTGCCAACTGCCATGTGGAAG
AGGTGCCTCCAGCTGCCCCGAGACGACCGGGCAGCCCATCGGCTATGGGAGGCCAGCAAGAGG
CTGGCAGGGCTTGGGCTGGGGAGGATGCTGAACCCGATGAAGACCCCCAGTCTGAGGACTC
AGAGGCCCATCTTCTAAGCACCCCCACCCCTGAGGAGCCACAGTTTCTCAACCTTACC
CCAGCCCTCAGAGCTCACCAGATTTGTCTAAGATGACGCACCGAATTCAGGCTAAAGTTGAG
CCTGAGATCCAGCTCTCCTTAACCCTCAGGCCAGGATGCTTGCCATGGCACTTCATGGTCCTT
GAAAACCTCGGATGTGTGTGAGGCCATGCCCTGGACACTGACGGGTTTGTGATCTTGACCTC
CGTGTTACTTTCTGGGGCCCCAAGCTGTGCCCTGGACATCTCTTTTCTGGTTGAAGGAAT
AATGGGTGATTATTTCTTCTGAGAGTGACAGTAACCCAGATGGAGAGATAGGGGTATGCT
AGACACTGTGCTTCTCGGAAATTTGGATGTAGTATTTTCAGGCCCCACCCCTATTGATTCTG
ATCAGCTCTGGAGCAGAGGCAGGGAGTTTGCAATGTGATGCACTGCCAACATTGAGAATTAG
TGAAGTATCCCTTTGCAACCGCTCTAGCTAGGTAGTTAAATACCCCCATGTTAATGAAGCG
GAATTAGGCTCCCGAGCTAAGGACTCGCCTAGGGTCTCACAGTGAGTAGGAGGAGGGCCTG
GGATCTGAACCCAGGGTCTGAGGCCAGGGCCGACTGCCGTAAAGATGGGTGCTGAGAAGTGA
GTGAGGGCAGGGCAGCTGGTATCGAGGTGCCCCATGGGAGTAAGGGGACGCCTTCCGGGCGG
ATGCAGGGCTGGGGTCATCTGTATCTGAAGCCCTCGGAATAAAGCGCGTTGACCGCCAAAA
AAAAAAAAAAAAAAAA

2019

<MW: 40849, pI: 7.98, NX(S/T): 0

amino acids 1-16

amino acids 46-49

amino acids 37-49 and 114-124

FIGURE 76

GGAGGAGACAGCCTCCTGGGGGGCAGGGGTTCCCTGCCTCTGCTGCTCCTGCTCATCATGGGAGGCAATGGCTCAG
 GACTCCCCGCGCCAGACTCCTAGTCCACCCCCAGGACCAAGCTGTGTCAGGGCCCTGGCCCTGCGAGGATGAGCTGC
 CAAGCTCTCAGGCCAGCCACCTCCACCAATCCGCTGGTTGCTGAATGGCAGCCCTGGAGCATGCTGCCCCAGAC
 CCACACCACTCTCTGCTGATGGGACCTTCTGCTGCTCAGCCCCCTGCCCGGGGACATGCCCCAGATGGCCAG
 GCCCTGTCCACAGACCTTGGGTGTCTACACATGTGAGGCCAGCAACCCGGCTTGGCAGCGGCAGTCAAGAGGGCGCT
 CGGCTGTCTGTGGCTGTCTCTCCGGGAGGATTTCAGATCCAGCTCCAGCTTGGGATCAGTGTGGCTGTGGTGGGTGAGCAG
 TTACTCTGGAATGTGGGGCGCCCTGGGGCCACCCAGAGCCCAAGTCTCATGGTGGAAAGATGGGAAACCCCTG
 GCCCTCCAGCCCGGAAGGCACACAGTGTCCGGGGGGTCCCTGCTGATGGCAGAGGACAGAGAAGATGACGAAGGG
 ACCTACATGTGTGTGGCCACCAACAGCGCAGGACATAGGGAGAGCCGCGCAGCCCGGGTTTCATCTCAGGAGCCCC
 CAGGACATACAGGAGCCTGTGGAGCTTCTGGCTGTGCGAATTCTAGCTGGAAATTTGACACTGCTGAACCCTGGAT
 CCTCGAGAGGGCCCCAAGCCTAGACCGGCGGTGTGGCTCAGCTGGAAGGTCTAGTGGCCCTGCTGCGCCTGCCCAA
 TCTTACACGGCCTTGTTCAGGACCCAGACTGCCCCGGGAGGCCAGGAGCTCCGTTGGGCAGAGGAGCTGCTGGCC
 GGCTGGCAGAGCGCAGAGCTTGGAGGCCCTCCACTGGGGCCAAAGACTACGAGTTCAAAGTGAGACCATCTCTTGGC
 CGGGCTCGAGGCCCTGACAGCAACGTGTCTCTCTGAGGCTGCCGGAAAAAGTCCCGAGCTGCCCACTCAGGAA
 GTGACTCTAAAGCCTGGCAATGGCACTGTCTTTGTGAGCTGGGTCCCAACCACTGCTGAAAACCAATGGCATC
 ATCCGTGGCTACAGGCTGTGGAGCTTGGGCAACACATCACTGCCACAGCCCACTGGACTGTAGTTGGTGAGCAG
 ACCCAGCTGGAATCGCCACCCATATGCCAGGCTCCTACTGCGTGCAGGTGGCTGCACTCTGGTCTGGAGCT
 GGGGAGCGCTAGACTGTCTGTGCTCTTCTTAGAGCAGGCCATGGAGCGACCCACCAAGAACCCAGTGAGCAT
 GGTCCCTGAGCCCTGGAGCAGCTGAGGGCTACCTTGAAGCGGCTGAGGTCAITGGCCACTGCGGTGTGCACT
 TGGCTGCTGCTTCTGGGCACCCGGTGTGTATCCACCCGCGCGCAGCTTAGGGTGTACCTTGGGGCCAGCTCTG
 TACAGATTAACCAAGTGAGGATGCCATCTTAAACACAGGATGATCAGACTGACTCCCAAGTGTGGCAGACAT
 TGGCGTTCCACCTCTGGCTCTCGGAGCTGAGCAGCAGCAGCCTCAGCAGTCCGTGGGGGCGGCGCCG
 GACCCACTAGACTGTCTGCTCTTGTCTCTCTGGGACTCCCGAAGCCCCGCGCTGCCCTGCTCTCAGACACC
 AGCACTTATATGCTCTCCTCATGCTGAGCTGCCCTCCAGTACCCAGCCAGGCTCCCAAGCTCCAGCT
 GTACAGCGCGTCCACCCAGCTGGCCAGCTCTCCAGCCCTGTGTTCCAGCTCAGACAGCCTCTGACGCCGAG
 GGACTCTTCTCTCCCGCTTGTCTCTGGCCCTCTCAGAGGCTTGGAGGCCAAAGAGACAGAGCTGCAGCAT
 GCCACAGTTCGCCCACTGTCTCCGGGCGAGCACTCTTGGAGCTCCGGGCTGTGAGTTAGGAAATAGAGGTTCC
 AAGAACCTTCCCAAAGCCAGGAGCTGTGCCCAAGCTCTGTTTGGCTGGCGGGCCCTGGGACCGAAACTCCTC
 AGCTCTCAAATGAGCTGGTTACTCGTCACTCTCCCTCAGCAGCCCTCTTCTCTCATGAACTCCCCCAACTCAG
 AGTCAACAGAGCCAGCTCCCGTGGCACCACAGGCTCCCTCTCATCTCTGCTGCCAGCAGCCCCCATCCCATC
 CTTAGCCCCCTGCACTCCCTCAGCCCCCAGGGCTCTTCCCTCTCTGCCCCCAGCCAGCTTCCAGTCCGCTGTCC
 AGCTCTCACTGTCTCTCTGGGGGAGGATCAAGACAGCGTGTGACCCCTGAGGAGGTAGCCCTGTGCTTGGAA
 CTCAGTGAGGGGTGAGGAGACTCCAGGAACAGCGTCTCTCCCATGCCAAGGGCTCTTCCACCCCAACCATAT
 GGGTACATCAGCGTCCCAACAGCTCAGAGTTACAGGACATGGGAGGAGCTGGAGGAGGGGTGGGGCCCAAGGGG
 GGGTCTTGCTGTGGCCACTCCGGCCCTGCTCACCCCCACCCAGCGAGGGCTCTTAGCCAAATGTTTGGGGC
 TCAGCCTCTGAGGACAATGCCGCGAGCGCCAGAGCCAGCTTGTGAGCTCTCCGATGGCTCTTCTCTCGCTGAT
 GCTCACTTTGGCCGGGCCCTGGCAGTGGCTGTGGATAGCTTTGGTTTCGGTCTAGAGCCCGAGGGAGGACATGCTG
 GTCTTTCATAGATGCTCATCACTCCCTCCCCACCGGATGAGATCTTCTGACCCCCAACCTCTCCCTCGCCCTG
 TGGAGGTGGAGGCCAGATGGTTGGAAGACATGGAGGTGAGCCACCCAGCGAGGCTCTTAGCCAAATGTTTGGGGC
 TGGCCCCCTGACTCTCAGATCTCTTCCAGAGAGATCTAGCTTCACTGTGCTATGCCCAAAGCTGGTGTCTCTCT
 GTAGATTACTCTCTGAACCGTGTCTCTGAGACTTCCAGCAGCGGAATCAGAAACACTTCTCTCTGTCACCACTCAG
 ACTCGGGCTGTGGTGTGGGGTCTTGGCCCTGTGTTTCTCTGCACTGGGGTCCACCTCTCCCAAGCCTCAGAGAG
 TTTCTCCCTCCAGATTTGTGAAACAATGAAACAAATATAGAGCAAGCTGACCTGGAGCCCTCAGGAGGCAAA
 ACATCATCTCCACCTGACTCCTAGCACTGCTTTCTCTCTGTGCCATCCACTCCCAACCAAGGTTGTTTGGC
 CTGAGGAGCAGCCCTGCTGCTCTTCCCAACCATTTGGATCAGGAAAGTGGAGGAGCAGAGGCTGCTTT
 GTGGAGGACAGCACTGGCTGTGGGAGAGGGCTGTGGGAGAGGAGCTTCTCGGAGCCCCCTCTCAGCCTTACTT
 GGGCCCTCTCTCTAGAGAAGAGCTCAACTCTCTCCCAACTCAGCATGGAAAGAAATTAATATGATGCCACT
 AGGCACCTGAGGCCCTACTCTATGCCAAACAAGGGTCAAGGCTGGGTCTAGCGAGGATGCTGAGAGAGGAGG
 TATGAGCCGTAGTCAAAAGCACCCTCTCGTACTGTGTTGCTACTATGAGCTTAAGAAATTTGATACCATATAAT
 GGTAAAAA

1001101.102401

FIGURE 77

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA41404

<subunit 1 of 1, 985 aa, 1 stop

<MW: 105336, pI: 6.55, NX(S/T): 7

MGGMAQDSPPQILVHPQDLFQGGPGPARMSCQASQPPTIRWLLNGQPLSMVFPDPHLLLP
DGTLLLLQPPARGHAHDGQALSTDLGVYTCEASNRLGTAVERGARLSVAVLREDFQIQPRDM
VAVVGEQFTLECGPPWGHPEPTVSWWKDGKPLALQPGRHTVSGGSLLMARAESDEGTYMCV
ATNSAGHRESRAARVSIQEPQDYTEPVELLAVRIQLENTLLNPDPAEGPKPRPAVWLSWKV
SGPAAPAQSYTALFRTQTAPGGQGAPWAEELLAGWQSAELGGLHWGQDYEFKVRPSSGRARG
PDSNVLLLRLEPKVPSAPPQEVTLKPGNGTVFVSWVPPPAENHNGIIRGYQVWSLGNSTSLPP
ANWTVVGEQTQLEIATHMPGSYCVQVAAVTGAGAGEPSRPVCLLLEQAMERATQEPSEHGPW
TLEQLRATLKRPEVIATCGVALWLLLLGTAVCIHRRRRARVHLGPGLYRYTSEDAILKHRMD
HSDSQWLADTWRSTSGSRDLSSSSSLSSRLGADARDPLDCRRSLLSWDSRSPGVPLLPDTST
FYGSLIAELPSSTPARPSFQVPAVRRLPFLAQLSSPSSSSDLSCSRRLSSPRLSLAPAEA
WKAKKKQELQHANSPLLGRGSHSLELRACELGNRGSKNLSQSPGAVPQALVAWRALGPKLLS
SSNELVTRHLFPAPLFPHPHETPTQSQQTQPPVAPQAPSSILLPAAPAPIILSPCSPPSPQASS
LSGSPASSRLSSSSSLSSLGEDQDSVLTPEEVALCLELSEGEETPRNSVSPMPRAPSPPTTY
GYISVPTASEFTDMGRTGGGVGPKGGVLLCPPRPCLTPTPSEGLANGWGSASEDNAASARA
SLVSSSDGSFLADAHFARALAVVDSFGFGLPREADCVFIDASSPSPRDEIFLTPNLSLP
LWEWRPDWLEDMEVSHTQRLGRGMPFWPPDSQISSQRSQLHCRMPKAGASFPVDYS

Important features:

Transmembrane domain:

amino acids 448-467

N-glycosylation sites:

amino acids 224-227, 338-341, 367-370, 374-377, 658-661 and 926-929

N-myristoylation sites.

amino acids 47-52, 80-85, 88-93, 99-104, 105-110, 181-186, 272-277, 290-295, 355-360, 403-408, 462-467, 561-566, 652-657, 849-854 and 876-881

Phosphotyrosine interaction domain proteins

amino acids 740-753

FIGURE 78

CTCCCACGGTGTCACGCGCCAGAAATGCGGCTTCTGGTCTGCTATGGGGTTGCCGTGCTGCT
 CCCAGGTTATGAAGCCCTGGAGGGCCAGAGGAAATCAGCGGGTTTGAAGGGGACACTGTGT
 CCCTGCAGTGCACCTACAGGGAGAGCTGAGGGACCACCGGAAGTCTGGTGACGGAAGGTT
 GGGATCCTCTTCTCTGCTGCTCTGGCACCATCTATGCAGAAGAAGGCAGGAGACAAT
 GAAGGGCAGGGGTGCCATCCGTGACAGCGCCAGGGCTCTCGCTCATTGTGACCCCTGTGGA
 ACCTCACCCCTGCAAGACGCTGGGGAGTACTGGTGTGGGGTCGAAAAACGGGGCCCGATGAG
 TCTTTTACTGATCTCTCTGTTCGTTCTTTCAGGACCCCTGCTGTCTCCCTCCCTTCTCCAC
 CTTCCAGCCTCTGGCTACAACACGCTGACGCCAAGGC AAAAGCTCAGCAAAACCCAGCCCC
 CAGGATTGACTTCTCTGGGCTTACCCGGCAGCCACCACAGCCAAGCAGGGGAAGACAGGG
 GCTGAGGCCCCCTCCATTGCCAGGGACTTCCAGTACGGGCACGAAAGGACTTCTCAGTACAC
 AGGAACCTCTCCTCACCCAGCGACCTCTCTCTGTCAGGGAGCTCCCGCCCCCATGTCAGC
 TGGACTCCACCTCAGCAGAGGACACCAGTCCAGCTCTCAGCAGTGGCAGCTTAAGCCAGG
 GTGTCCATCCCGATGGTCCGATACCTGGCCCCAGTCTGGTCTGCTGAGCCTTCTGTGAGC
 CGCAGGCCTGATCGCCTTCTGCAGCCACCTGCTCCTGTGGAGAAAGGAAAGCTCAACAGGCCA
 CGGAGACACAGAGGAACGAAAGTTCTGGCTCTCAGCTTGACTGCGGAGAAAGGAAGCC
 CCTTCCAGGCCCCCTGAGGGGACGTGATCTCGATGCCTCCCTCCACACATCTGAGGAGGA
 GCTGGGCTTCTCGAAGTTTGTCTCAGCGTAGGGCAGGAGGCCCTCTGGCCAGGCCAGCAGT
 GAAGCAGTATGGCTGGGTGATCAGCACCGGATTCGCCAAAGCTTTCCACCTCAGCCTCAGAG
 TCAGCTGCCCGGACTCAGAGGCTCTCCCAACCTCCCCAGGCTCTCCTCTGTGATGTTCCA
 GCCTGACCTAGAAGCGTTTGTGAGCCCTGGAGCCAGAGCGGTGCCCTTGTCTTCCGGCTG
 GAGACTGGGACATCCCTGATAGGTTACATCCCTGGGCAGAGTCAACAGGCTGCTGACCCCTCA
 GCAGGGCCAGACAAGGCTCAGTGGATCTGGTCTGAGTTTCAATCTGCCAGGAACCTCCTGGGC
 CTCATGCCCCAGTGTGCGACCCCTGCCCTTCCCTCCCACTCCAGACCCACCTTGTCTTCCCTCCC
 TGGCGTCTCAGACTTAGTCCCACGGTCTCCTGCATCAGCTGGTGTGTAAGAGGAGCATGCT
 GGGGTGAGACTGGGATTCCTGGCTTCTCTTTGAACCACTGCATCCAGCCCTTCAGGAAGCCT
 GTGAAAAACGTGATTCCTGGCCCCCAAGACCCCAAAACCATCTCTGGGCTTGGTGCAG
 GACTCTGAATTTCAACAATGCCAGTGACTGTGCACTTGAGTTTGAGGGCCAGTGGGCCCTG
 ATGAACGCTCAACCCCTTCAGCTTAGAGTCTGCATTTGGGCTGCTGAGCTCTCCACCTGCC
 CAATAGATCTGCTCTGTCTGCGACACCAGATCCACGTGGGACTCCCTGAGGCCTGCTAAG
 TCCAGGCCCTTGGTCAGGTCAGGTGCACATTGCAGGATAAGGCCAGGACCCGCCACAGAAGTGG
 TTGCCCTTNCACATTGGCCCTCCCTGGNCCATGCCTTCTGGCTTTGGAAAAAATGATGAAGA
 AAACCTTGGCTCCTTCTTGTCTGGAAGGGTTACTTGCTATGGGTTCTGGTGGCTAGAGA
 GAAAAGTAGAAAAACAGATGCACTAGGTGTCTAACACAGAGGAGAGTAGGAACAGGGCGG
 ATACCTGAAGTGACTCCGAGTCCAGGCCCTGGAGAGGGGTTCGGGGGTGGTGGGTAAGTA
 GCACAACACTATTTTTTTTCTTTTCCATATTATTGTTTTTAAAGACAGAATCTCGTGCT
 GCTGCCCGGCTGGAGTGCAGTGGCAGCATCTGCAAACTCCGCTCTCTGGGTTCAAGTGATT
 CTTCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGCACGCACCACCACCTGGCTAATT
 TTTGTACTTTTAGTAGAGATGGGGTTTACCATTGTTGGCCAGGCTGGTCTTGAACCTCTGCAC
 CTCAAATGAGCCTCCTGCTTTCAGTCTCCCAAATTGCCGGGATGACGGCATGAGCCACTGTG
 TCTGGCCCTATTTCTTTAAAAAGTGAATAAGAGTTGTTAGTATGCAAAACTTGGAAAG
 ATGGAGGAGAAAAAGAAAGGAAGAAAAAATGTCACCATAGCTCACCAGAGACTCATAT
 TATTTCGTTTTGTTGACTTCTTCCACTTTTCTTCTGACATAATTTGCCGGTGTCTCTT
 TTTACAGAGCAATTTATCTGTATATACAACCTTTGTATCTCGCTTTTCCACCTTTATCGTTCC
 ATCATTTATTTCAGCACTTCTCTGTGTTTTACAGACCTTTTATAAATAAAATGTTTCATCA
 GCTGCATAAAAAAAAAAAAAA

FIGURE 79

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44196
<subunit 1 of 1, 332 aa, 1 stop
<MW: 36143, pI: 5.89, NX(S/T): 1
MRLLVLLWGCLLLPGYEALGPEEISGFEGDTVSLQCTYREELRDHRKYWCRKGGILFSRCS
GTIYABEEGQETMKGRVSIIRDSRQELSLIVTLWNLTLDAGEYWCVEKRGPDSELLISLFV
FPGPCPPSPSPFTFQPLATTRLQPKAKAQQTQPPGLTSPGLYPAATTAKQKGTGAEAPPLPG
TSQYGHERTSQYTGTSPHPATSPPPAGSSRPPMQLDSTSAEDTSPALSSGSSKPRVSIIPMVR I
LAPVLVLLSLLSAAGLIAFCSHLLLRKEAQQATETQRNEKFWLSRLTAEKEAPSQAPEGD
VISMPPLHTSEELGFSKFVSA

Important features:

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 248-269

N-glycosylation site.

amino acids 96-99

Fibrinogen beta and gamma chains C-terminal domain.

amino acids 104-113

Ig like V-type domain:

amino acids 13-128

Figure 1. Schematic representation of the experimental design. The subjects were divided into two groups: the control group (CG) and the experimental group (EG). The CG was divided into two subgroups: the control group (CG) and the control group (CG). The EG was divided into two subgroups: the experimental group (EG) and the experimental group (EG). The subjects were divided into two groups: the control group (CG) and the experimental group (EG). The CG was divided into two subgroups: the control group (CG) and the control group (CG). The EG was divided into two subgroups: the experimental group (EG) and the experimental group (EG).

TTGTGACTAAAAGCTGGCCTAGCAGGCCAGGGAGTGCAGCTGCAGGCGTGGGGGTGGCAGGA
GCCGCAGAGCCAGAGCAGACAGCCGAGAAACAGGTGGACAGTGTGAAAGAACCAGTGGTCTC
GCTCTGTTGCCCAGGCTAGAGTGTACTGGCGTGATCATAGCTCACTGCAGCCTCAGACTCCT
GGACTTGAGAAATCCTCTGCCTTAGCCTCCTGCATATCTGGGACTCCAGGGGTGCACTCAA
GCCCTGTTTCTTCTCCTTCTGTGAGTGGACCACGGAGGCTGGTGAAGCTGCCTGTATCCCAA
AGCTCAGCTCTGAGCCAGAGTGGTGGTGGCTCCACCTCTGCCGCCGCATAGAAGCCAGGAG
CAGGGCTCTCAGAAGCGGGTGGTGCCACAGCTGGGATCATGTTGTTGGCCCTGGTCTGTCTGC
TCAGCTGCCTGCTACCTCCAGTGAGGCCAAGCTCTACGGTCGTTGTGAAGTGGCCAGAGTG
CTACATGACTTCGGGCTGGACGGATACCGGGGATACAGCCTGGCTGACTGGTCTGCCTTGC
TTATTTCAACAAGCGGTTTCAACGCAGCTGCTTTGGACTACGAGGCTGATGGGAGCACCAACA
ACGGGATCTTCAGATCAACAGCCGGAGGTGGTGAGCAACCTCACCCGAACGTCCCCAAC
GTGTGCCGGATGTACTGCTCAGATTTGTTGAATCCTAATCTCAAGGATACCGTTATCTGTGC
CATGAAGATAACCCAAGAGCCTCAGGGTCTGGGTTACTGGGAGGCCTGGAGGCATCACTGCC
AGGAAAAAGACCTCACTGAATGGGTGGATGGCTGTGACTTCTTAGGATGGACGGAACCATGCA
CAGCAGGCTGGGAAATGTGGTTTGGTTTCCTGACCTAGGCTTGGGAAGACAAGCCAGCAGAATA
AAGGATGGTTGAACGTGAAA

FIGURE 81

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52187

<subunit 1 of 1, 146 aa, 1 stop

<MW: 16430, pI: 5.05, NX(S/T): 1

MLLALVCLLSCLLPSEAKLYGRCELARVLHDFGLDGYRGYSLADWVCLAYFTSGFNAAALD
YEADGSTNNGIFQINSRRWCNSLTPNVPNVCRMYSDDLNPNLKDTVICAMKITQEPQGLGY
WEAWRHHHCQKDLTEWVDGCDF

Important features:

Signal peptide:

amino acids 1-18

N-myristoylation site.

amino acids 67-72

Homologous region to Alpha-lactalbumin / lysozyme C proteins.

amino acids 34-58 (catalytic domain), 111-132 and 66-107

FIGURE 82

AGCCGCTGCCCCGGGCCGGGCGCCCGGGCGGCACCATGAGTCCCCGCTCGTGCTGCGTTCC
GCTGCGCCTCCTCGTCTTCGCGCTCTTCTCAGCCGCGCGAGCAACTGGCTGTACCTGGCCA
AGCTGTGTCGTCGGTGGGAGCATCTCAGAGGAGGAGACGTGCCAGAACTCAAGGGCCTGATC
CAGAGGCAGTGCAGATGTGCAAGCGGAACCTGGAAGTCATGGACTCGGTGCGCCGCGGTGC
CCAGCTGGCCATTGAGGAGTGCCAGTACCAGTTCGGAACCGGCGCTGGAACCTGCTCCACAC
TCGACTCCTTGCCCGTCTTCGGCAAGGTGGTGACGCAAGGGACTCGGAGGCGGCCTTCGTG
TACGCCATCTCTTCGGCAGGTGTGGCCTTTGCAGTGACGCGGCGTGCAGCAGTGGGGAGCT
GGAGAAGTGCGGCTGTGACAGGACAGTGACATGGGGTCAAGCCACAGGGCTTCAGTGGTCAG
GATGCTCTGACAACATCGCCTACGGTGTGGCCTTCTCAGTCTGTTTGTGGATGTGCGGGAG
AGAAGCAAGGGGGCTCGTCCAGCAGAGCCCTCATGAACCTCCACAACATGAGGCCGCGAG
GAAGGCCATCTGACACACATGCGGGTGGAAATGCAAGTGCCACGGGGTGTGAGGCTCCTGTG
AGGTAAAGACGTGCTGGCGAGCCGTGCCGCCCTTCCGCCAGGTGGGTGACGCACTGAAGGAG
AAGTTTGATGTTGCCACTGAGGTGGAGCCACGCCGCTGCGGCTCTCCAGGGCACTGGTACC
ACGCAACGACAGTTCAAGCCGACACAGATGAGGACCTGGTGTACTTGGAGCCTAGCCCCG
ACTTCTGTGAGCAGGACATGCGCAGCGCGTGTGCGGCACGAGGGGCCGACATGCAACAAG
ACGTCCAAGGCCATCGACGGCTGTGAGCTGCTGTGCTGTGGCCGCGCTTCCACACGGCGCA
GGTGGAGCTGGCTGAACGCTGCAGCTGCAAAATCCACTGGTGTGCTTCGTCAAGTGCCGGC
AGTGCCAGCGGCTCGTGGAGTTGCACACGTGCCGATGACCGCCTGCCTAGCCCTGCGCCGGC
AACCACCTAGTGGCCAGGGAAGGCCGATAATTTAAACAGTCTCCACCACCTACCCCAAGA
GATACTGGTTGTATTTTGTCTGTTTGGTTTGGTCTCCTCATGTTATTTATTGCCGAA
ACCAGGCAGGCAACCCCAAGGGCACCAACCAGGGCCTCCCAAAGCCTGGGCTTTGTGGCT
GCCACTGACCAAGGGACCTTGCTCGTGCCGCTGGCTGCCCGCATGTGGCTGCCACTGACCA
CTCAGTTGTTATCTGTGTCGCTTTTCTACTTGACAGCTAAGGTGGAGTAACAAGGAGTAT
TACCACCACATGGCTACTGACCGTGTGTCATCGGGGAAGAGGGGGCCTTATGGCAGGGAATA
GGTACCGACTTGATGGAAGTCACACCCTCTGAAAAAAGAACTTTAACTCTCCAGCACACA
TACACATGGACTCTGGCAGCTTGAGCCTAGAAGCCATGTCTCTCAATGCCCCTGAGAAAGG
GAACAAGCAGATACAGGTCAAGGGCACCAGGTTATTTAGCCCTTACATGGACAGCTAGA
GGTTCGATATCTGTGGTCTTCCAGGCAAGAAGAGGAGATGAGAGCAAGAGACGACTGAA
GTCCCAACCTTAGAACCCAGCCCTGCCCGAGCCTGCCCTGGGAAGAGGAACTTAACCACTCC
CCAGACCCACCTAGGCAGGCATATAGGCTGCCATCCTGGACAGGGATCCCGGCTGTGCCTT
TGCAGTCATGCCCGAGTCACCTTTCACAGCGCTGTTCTCCATGAAACTGAAAAACACACAC
ACCTGCGAGA
GAGAGGGAGGAAAGGGCTGTGCCTTTGCAGTCATGCCGAGTCACCTTTCACAGCACTGTTCTCT

FIGURE 83

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48328

<subunit 1 of 1, 351 aa, 1 stop

<MW: 39052, pI: 8.97, NX(S/T): 2

MSPRSCRLRSLRLLVFAVFSAAASNWLYLAKLSSVGSISEEETCEKLGKLIQRQVQMCKRNLE
VMDSVRRGAQLAIEECQYQFRNRRWNCSTLDSLVPFGKVVTQGTREAAFVYAISSAGVAFV
TRACSSGELEKCGCDRTVHGVSPQGFQWSGCSNLIAYGVAFSQSFVDVRERSKGASSSRALM
NLHNNEAGRKAILTHMRVECKCHGVSGSCEVKTWRAVPPFRQVGHALKEKFDGATEVEPRR
VGSSRALVPRNAQFKPHTDEDLVYLEPSPDFCEQDMRSGVLGTRGRTCNKTSKAIDGCELLC
CGRGFHTAQVELAERCCKFHWCCFVKCRQCQRLVELHTCR

Important features:

Signal peptide:

amino acids 1-22

N-glycosylation sites.

amino acids 88-91 and 297-300

Wnt-1 family signature.

amino acids 206-215

Homologous region to Wnt-1 family proteins

amino acids 183-235, 305-350, 97-138, 53-92 and 150 -174

FIGURE 84

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCTGGGTGCCTGCAT
CGCCATGGACACCACCAGGTACAGCAAGTGGGGCGGACGCTCCGAGGAGTCCCCGAGGGC
CCTGGGACGCTGGGTGCACTGGAGCAGGAGACCCCTCTTCTTGGCCCTGGCTGTCTGGTC
ACCAAGTCTCTTTGGGCTGTGATTCTGAGTATCCTATTGTCCAAGGCCTCCACGGAGCGCG
GGCGCTGCTTGACGGCCACGACCTGCTGAGGACAAACGCCTCGAAGCAGACGGCGGCGCTGG
GTGCCCTGAAGGAGGAGGTCGAGAGCTGCCACAGCTGTGCTCGGGGACGAGCGCGCAGCTG
CAGACCACGCGCGCGGAGCTTGGGGAGGCGCAGGCGAAGCTGATGGAGCAGGAGAGCGCCCT
GCGGGAACTGCGTGAGCGCGTGACCCAGGGCTTGGCTGAAGCCGGCAGGGGCCGTGAGGACG
TCCGCACTGAGCTGTTCCGGGCGCTGGAGGCCGTGAGGCTCCAGAACAACCTCTGCGAGCCG
TGCCCCACGTCGTGGCTGTCTTCGAGGGCTCCTGCTACTTTTTCTCTGTGCCAAAGACGAC
GTGGGCGGCGGCGCAGGATCACTGCGCAGATGCCAGCGCGCACCTGGTGATCGTTGGGGGCC
TGGATGAGCAGGGCTTCTCTCACTCGGAACACGCGTGGCCGTGGTTACTGGCTGGGCCTGAGG
GCTGTGCGCCATCTGGGCAAGGTTCAAGGGCTACCAAGTGGGTGGACGGAGTCTCTCTCAGCTT
CAGCCACTGGAACAGGGAGAGCCCAATGACGCTTGGGGGCGCGAGAACTGTGTATGATGC
TGCACACGGGGCTGTGGAACGACGCCCGTGTGACAGCGAGAAGGACGGCTGGATCTGTGAG
AAAAGGCACAACCTGCTTGACCCCGCCAGTCCCTGGAGCCGCGCCCATTCAGCATGTCTGTA
TCCTGGGGGCTGTCTACCTCCCTGGCTCCTGGAGCTGATTGCCAAAGAGTTTTTTCTCTCCT
CATCCACCGCTGTCTGAGTCTCAGAAACACTTGGCCCAACATAGCCCTGTCCAGCCCAGTGCC
TGGGCTCTGGGACCTCCATGCCGACCTCATCCTAACTCCACTCACGAGACCCAACTAACC
TCCACTAGCTCCAAATCCCTGCTCCTGCGTCCCCGTGATATGCCCTCCACTTCTCTCCCTAA
CCAAGTTAGGTGACTGAGGACTGGAGCTGTTTGGTTTCTCGCATTTTCCACCAAACTGGA
AGCTGTTTTTGACGCTGAGGAAGCATCAATAAATATTGAGAAATGAAAAA

FIGURE 85

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56352

<subunit 1 of 1, 293 aa, 1 stop

<MW: 32562, pI: 6.53, NX(S/T): 2

MDTTRYSKWGGSSSEEVPGGPWGRVHWSRRPLFLALAVLVTTVLWAVILSILLSKASTERAA
LLDGHDLLRTNASKQTAALGALKKEEVGDCHSCCSGTQAQLQTTRAELGEAQAKLMEQESALR
ELRERVTOGLAEAGRGREDVRTELFRALAEAVRLQNNSCPCPTSWLSFEGSCYFFSVPKTTW
AAAQDHCADASAHLVIVGGLDEQGFLTRNTRGRGYWLGLRAVRHLGKVQGYQWVDGVSLSFS
HWNQGEFNDAWGRENCVMMLHTGLWNDAPCDSEKDGWICEKRHNC

Important features:

Type II transmembrane domain:

amino acids 31-54

N-glycosylation sites.

amino acids 73-76 and 159-162

Leucine zipper pattern.

amino acids 102-123

N-myristoylation sites.

amino acids 18-23, 133-138 and 242-247

C-type lectin domain signature.

amino acids 264-287

FIGURE 87

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53971

><subunit 1 of 1, 331 aa, 1 stop

><MW: 35844, pI: 5.45, NX(S/T): 2

MENPSPAAALGKALCALLLATLGAAGQPLGGESICSARAPAKYSITFTGKWSQTAFPKQYPL
FRPPAQWSSLLGAHSSDYSMWRKNQYVSNGLRDFAEERGEAWALMKEIEAAGEALQSVHEVF
SAPAVPSGTGQTSAELEVQRRHSLVSFVVRIVPSPDWFVGVDSLDLCGDRWREQAALDLYP
YDAGTDSGFTFSSPNFATIPQDTVTEITSSSPSPHPANSFYYPRLKALPPIARVTLRLRQSP
RAFIPAPVLPSPRDNEIVDSASVPETPLDCEVSLWSSWGLCGGHCGR LGTKSRTRYVRVQPA
NNGSPCPELEEEAECPDNCV

Important features:

Signal peptide:

amino acids 1-26

FIGURE 88

GGCGGCGTCCGTGAGGGGCTCCTTTGGGCAGGGGTAGTGTTTGGTGTCCCTGTCTTGCGTGGA
TATTGACAAACTGAAGCTTTTCTGCACCACTGGACTTAAGGAAGAGTGACTCGTAGGCGGA
CAGCTTTAGTGGCCGGCCGGCCGCTCTCATCCCCGTAAGGAGCAGAGTCCTTTGTACTGAC
CAAGATGAGCAACATCTACATCCAGGAGCCTCCCAGAAATGGGAAGGTTTTATTGAAAACTA
CAGCTGGAGATATTGACATAGAGTTGTGGTCCAAAGAAGCTCCTAAAGCTTGCAGAAATTTT
ATCCAACCTTTGTTTGGGAAGCTTTATTATGACAATACCATTTTTCATAGAGTTGTGCCTGGTTT
CATAGTCCAAGGCGGAGATCCTACTGGCACAGGGAGTGGTGGAGAGTCTATCTATGGAGGCG
CATTCAAGAGTGAATTTTCATTACGGTTGCGTTTTTAATCGGAGAGGACTGGTTGCCATGGCA
AATGCTGGTTCTCATGATAATGGCAGCCAGTTTTTCTTACACTGGGTCGAGCAGATGAACT
TAACAATAAGCATACCATCTTTGGAAAGGTTACAGGGGATACAGTATATAACATGTTGCGAC
TGTCAGAAGTAGACATTGATGATGACGAAAGACCATAATCCACACAAAAATAAAAGCTGT
GAGGTTTTGTTTAATCCTTTTGATGACATCATTCCAAGGGAATTAAGGCTGAAAAAAGA
GAAACCAGAGGAGGAAGTAAAGAAATTGAAACCCAAAGGCACAAAAATTTTAGTTTACTTT
CATTTGGAGAGGAAGCTGAGGAAGAAGAGGAGGAAGTAAATCGAGTTAGTCAGAGCATGAAG
GGCAAAAGCAAAAGTAGTCATGACTTGCTTAAGGATGATCCACATCTCAGTTCTGTTCCAGT
TGTAAGAAAGTAAAAAGGTGATGCACCAGATTTAGTTGATGATGGAGAAGATGAAAGTGCAG
AGCATGATGAATATATTGATGGTGATGAAAAGAACCCTGATGAGAGAAAGAAATGCCAAAAAA
TTAAAAAAGGACACAAGTGCGAATGTTAAATCAGCTGGAGAAGGAGAAGTGGAGAAGAAATC
AGTCAGCCGAGTGAAGAGCTCAGAAAAAGCAAGACAATTAACCGGGAACCTTTAGCAG
CAAAAACAAAAAAGTAGAAAAATGCAGCAAAACAAGCAGAAAAAAGAAGTGAAGAGGAAGAA
GCCCTCCAGATGGTGCTGTTGCCGAATACAGAAGAGAAAAGCAAAAGTATGAAGCTTTGAG
GAAGCAACAGTCAAAGAAGGGAACCTCCCGGAAGATCAGACCCTTGCACTGCTGAACCAGT
TTAAATCTAAACTCACTCAAGCAATTGCTGAAACACCTGAAAAATGACATTCCTGAAACAGAA
GTAGAAGATGATGAAGGATGGATGTCACATGTACTTCAGTTTGAGGATAAAAGCAGAAAGT
GAAAGATGCAAGCATGCAAGACTCAGATACATTTGAAATCTATGATCCTCGGAATCCAGTGA
ATAAAAGAAGGAGGAAGAAAGCAAAAGCTGATGAGAGAGAAAAAGAAAGAAGATAAAAAT
GAGAATAATGATAACCAGAACTTGCTGGAATGTGCCTACAATGGCCTTGTAACAGCCATTG
TTCCCAACAGCATCACTTAGGGGTGTGAAAGAAGTATTTTTGAACCTGTTGCTCTGGTTTTG
AAAAACAATTATCTTTGTTTGGCAAATTGTGGAATGATGTAAGCAAAATGCTTTTGTTTACTGG
TACATGTGTTTTTCTTAGCTGACCTTTTATATTGCTAAATCTGAAATAAAATAACTTTCTCT
TCCACAAAAA

FIGURE 89

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50919
><subunit 1 of 1, 472 aa, 1 stop
><MW: 53847, pI: 5.75, NX(S/T): 2
MSNIYIQEPPTNGKVLLKTTAGDIDIELWSKEAPKACRNFIQLCLEAYYDNTIFHRVVPGFI
VQGGDPTGTGSGGESIYGAPFKDEFHSRLRFNRRGLVAMANAGSHDNGSQFFFTLGRADELN
NKHTIFGKVTGDTVYNMLRLSEVDIDDDERPHNPHKIKSCEVLFPFDDIIPREIKRLKKEK
PEEEVKKLKPKGTKNFSLLSFGEEAEEEEEEVNVRVSQSMKGKSKSSHDLKDDPHLSSVPV
ESEKGDAPDLVDDGEDESAEHDEYIDGDEKNLMRERIAKKLKKDTSANVKSAGEGEVEKKS
SRSEELRKEARQLKRELLAAKQKKVENAAKQAEKRSEEEAAPDGAVA EYRREKQKYEALRK
QQSKKGTSREDQTLALLNQFSKLTQAIAETPENDIPETEVEDDEGWMShVLQFEDKSRKVK
DASMQDSDTFEYDPRNPVNKRREESKKLMREKKERR
```

Important features:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 109-112 and 201-204

Cyclophilin-type peptidyl-prolyl cis-trans isomerase signature.

amino acids 49-66

Homologous region to Cyclophilin-type peptidyl-prolyl cis-trans isomerase

amino acids 96-140, 49-89 and 22-51

FIGURE 91

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44179
<subunit 1 of 1, 545 aa, 1 stop
<MW: 58934, pI: 9.45, NX(S/T): 4
MSTGFSFGSGTLGSTTVAAGGTSFTGGVFSFGTGTSSNPSVGLNFGNLGSTSTPATTSAPSSG
FGTGLFGSKPATGFTLGGTNTGALHTKRPQVVTKYGTQGGQMHVGKTPIQVFLGVFPFSRPP
LGILRFAPPEPPEPWKGIRDATTYPPGWSLALSPGWSAVARSRLTATSASRVQASLLPQPLS
VWGYRCLQESWQLASMYVSTRETRYKWLRFSEDCLYLNVYAPARAPGDPQLPVMVWFPGGAF
IVGAASSYEGSDLAAREKVVLVFLQHRLGIFGFLSTDDSHARGNWGLLDQMAALRWVQENIA
AFGGDPGNVTLFGQSAGAMSISGLMMSPLASGLFHRAISQSGTALFRLFITSNPLKVAKKVA
HLAGCNHNSTQILVNCLRALSGTKVMRVSNKMRFLQLNFORDPPEEIIWSMSPVVDGVVIPPDD
PLVLLTQGVSSVPYLLGVNNLEFNWLLPYNITKEQVPLVVEEYLDNVNEHDWKMLRNRMMMD
IVQDATFVYATLQTAHYHRETPMMGICPAGHATTRMKSTCSWILPQEWA

Important features:

Signal peptide:

amino acids 1-29

Carboxylesterases type-B serine active site.

amino acids 312-327

Carboxylesterases type-B signature 2.

amino acids 218-228

N-glycosylation sites.

amino acids 318-321, 380-383 and 465-468

FIGURE 93

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA54002
><subunit 1 of 1, 544 aa, 1 stop
><MW: 60268, pI: 9.53, NX(S/T): 3
MLLPLLSSLLGGSQAMDGRFWIRVQESVMVPEGLCISVPCSFYPRQDWTGSTPAYGYWFK
AVTETTTKGAPVATNHQSREVEMSTRGRFQLTGDPKAGNCSLVIRDAQMQDESQYFFRVERGS
YVTYNFMNDGFFLKVTLSFTPRPQDHNTDLTCHVD FSRKGVSAQRTVRLRVAYAPRDLVIS
ISRDNTPALEPQPQGNVPYLEAQKGQFLRLLLCAADSQPPATLSWVLQNRVLSSSHFWGPRPL
GLELPVGKAGDSGRYTCRAENRLGSGQQRALDLSVQYPPENLRVMVSQANRTVLENLGNGTSL
PVLEGQSLCLVCVTHSSPPARLSWTQRGQVLSPSQSPDPGVLELPRVQVEHEGEFTCHARHP
LGSQHVSLSLSVHYKKGLISTAFSNGAFLGIGITALLFLCLALIIMKILPKRRTQTETPRPR
FSRHSTILDYINVVPTAGPLAQKRNQKATPNSPRTPPPPGAPSPESKKNQKKQYQLPSFPEP
KSSTQAPESQESQEELHYATLNFPGVRPRPEARMPKGTQADYAEVKFQ

Important features:

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 399-418

N-glycosylation site.

amino acids 100-103, 297-300 and 306-309

**Immunoglobulins and major histocompatibility complex proteins
signature.**

amino acids 365-371

Index

TTAAA

FIGURE 95

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53906
><subunit 1 of 1, 772 aa, 1 stop
><MW: 87002, pI: 4.64, NX(S/T): 8
MNCYLLLRFMLGIPLLWPCLGATENSQTKVKQPVRSHLRVKRGWVWNQFFVPEEMNTTSHH
IGQLRSDLDNNGNNSFYQYKLLGAGAGSTFIIDERTGDIYAIQKLDREERSLYILRAQVIDIAT
GRAVEPESEFVIKVS DINDNEPKFLDEPYEAIVPEMSPEGTLVIQVTASDADDPSSGNNARL
LYSLLQGQPYFSVEPTTG VIRISSKMDRELQDEYWVIIQAKDMIGQPGALSGTTSVLIKLS D
VNDNKP IFKESLYRLTVSE SAPGTGTSIGTIMAYDNDIGENAEMDYSIEEDDSQTFDIITNHE
TQEGIVILKKKVD FEHQNHYGIRAKVKNNHVPQLMKYHTEASTTFIKIQVEDVDEPPLFL L
PYYVFEVFEETPQGSFVGVSATDPDNRKSPIRYSITRSKVFNINDNGTITTSNSLDREISA
WYNLSITATEKYNIEQISSIPLYVQVLNINDHAPEFSQYYETYVCENAGSGQVIQTISAVDR
DESIEEHHFYFNLSVEDTNNSSFIIIDNQDNTAVILTNRTGFNLQEEFVFIISILIADNGIP
SLTSTNTLTIHVCD CGDSGSTQTCQYQELVLSMGFKTEVIIAILICIMIIFGFI FLTLGLKQ
RRKQILFPEKSEDFRENIFQYDDEGGGEEDTEAFDIAELRSSITIMRERKTRKTTSAEIRSLY
RQSLQVGPD SAIFRKFILKLEANTDPCAPPFDSLQTYAFEGTGLAGSLSSLESASVSDQD
ESYDYLNELGPRFKRLACMFGSAVQSNN
```

Important features:

Signal peptide:

amino acids 1-21

Transmembrane domain:

amino acids 597-617

N-glycosylation sites.

amino acids 57-60, 74-77, 419-423, 437-440, 508-511, 515-518,
516-519 and 534-537

Cadherins extracellular repeated domain signature.

amino acids 136-146 and 244-254

FIGURE 96

ATTTCAAGGCCAGCCATATTTTTNTGTTGAACCAACAACAGGAGTCATAAGAATATTTTNTA
AAATGGATAGAGAACTGCAAGATGAGTATTGGGTAAATCATTCAAGCCAAGGACATGATTGGT
CAGCCAGGAGCGTTGTNTGGAACAACAAGTGATTAATTAACTTTTCAGATGTTAATGACAA
TAAGCCTATATTTAAAGAAAGTTTATACCGCTTGACTGTNTNTGAATCTGCACCCACTGGGA
NTTNTATAGGAACAATCATGGCATATGATAATGACATAGGAGAGAATGCAGAAATGGATTAC
AGCATTGAAGAGGATGATTCGCAAACATTTGACATTATT

10017191.102401

FIGURE 97

GCAACCTCAGCTTCTAGTATCCAGACTCCAGCGCCGCCCGGGCGCGGACCCCAACCCCGAC
 CCAGAGCTTCTCCAGCGCGCGCGCAGCGAGCAGGGCTCCCCGCCTTAACCTTCTCCGCGGGG
 CCCAGCCACCTTCGGGAGTCCGGGTGCCACCTGCAAACTCTCCGGCTTCTGCACCTGCCA
 CCCCAGGACCGCGCGGCCCGAGCGAGTCATGCGCCAACGCGGGCTGCAGCTGTGTGGG
 TTCATTCTCGCCTTCTCGGGATGGATCGCGGCCATCGTCAGACTGCCCTGCCCCAGTGGAG
 GATTACTCCTATGCGCGGCACAACATCGTGACCGCCAGGCCATGTACGAGGGGCTGTGGA
 TGTCTCGGTGTCGCGAGAGCACCGGGCAGATCCAGTGCAAACTCTTGACTCCTTGCTGAAT
 CTGAGCAGCACATTGCAAGCAACCCGTGCCTTGATGGTGGTTGGCATCCTCTCGGAGTGAT
 AGCAATCTTTGTGGCCACCGTTGGCATGAAGTGATGAAGTGCTTGAAGACGATGAGGTGC
 AGAAGATGAGGATGGCTGTCAATTGGGGGTGCGATATTCTTCTTGCAGGTCTGGCTATTTTA
 GTTGCCACAGCATGGTATGGCAATAGAATCGTTCAAGAATCTATGACCCATGACCCCACT
 CAATGCCAGGTACGAATTTGGTCAGGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTCGCC
 TTCTGGGAGGTGCCCTACTTTGTGTTCTGTGCCGAAAAACAACCTTACCACAACCA
 AGGCCCTATCAAAAACCTGCACCTTCCAGCGGGAAAGACTACGTTGACACAGAGGCAAAAG
 GAGAAATCATCTTGAACAACCAACGAAAATGGACATTGAGATACATTAACTATTAGGAC
 CTTAGAATTTTGGGTATTGTAATCTGAAGTATGGTATTACAAAAACAAACAAACAAAA
 ACCCATGTGTTAAAACTCAGTGCTAAACATGGCTTAATCTTATTTATCTTCTTCTCA
 ATATAGGAGGGGAAGATTTTCCATTGTATTACTGCTTCCCATGGAGTAATCATACTCAAAT
 GGGGGAAGGGGTGCTCCTTAAATATATATAGATATATATACATGTTTTCTATTAAAA
 ATAGACGTAATAATACTATTCTCATTATGTTGATACATAGCATATAAATATCTTAAAT
 AGGTAAATGTATTTAATTTCCATATTGATGAAGATGTTTATTGGTATATTTCTTTTCTGCC
 TTATATACATATGTAACAGTCAAATATCATTTACTCTTCTTATTAGCTTTGGGTGCCCTTG
 CCACAAGACCTAGCCTAATTTACCAAGGATGAATCTTTCAATTCTTATGCGTGCCCTTTT
 CATATACTATTTTATTTTTTACCATAATCTTATAGCACTTGCATCGTTATTAAGCCCTTAT
 TTGTTTGTGTTTCATTGGTCTCTATCTCCTGAATCTAACACATTTTATAGCCTACATTTTA
 GTTTCTAAAGCCAAGAAGATTTATTACAAATCAGAACTTTGGAGGCAAACTTTCTGCAATG
 ACCAAAGTGATAAATTCCTGTTGACCTTCCACACAATCCCTGACTCTGACCCATAGCACT
 CTTGTTGCTTTGAAAATATTTGTCCAATTGAGTAGCTGCATGCTGTTCCCCCAGGTGTGT
 AACACAACCTTTATTGATTGAATTTTAAAGCTACTTATTATAGTTTTTATATCCCCCTAACT
 ACCTTTTGTGTTCCCATCTCCTAATTTGATTGTTTCCCAAGTGTAATATCATGCGTTTAT
 TATCTTCCTAATAAGGTGTGGTCTGTTTGTCTGAACAAAGTGCTAGACTTTCTGGAGTGATA
 ATCTGGTGACAAATATTTCTCTGTAGCTGTAAGCAAGTCACTTAATCTTCTACCTCTTTT
 TTCATCTGCCAAATGAGATAATGATACTTAAACAGTTAGAAGAGGTAGTGTGAATATTA
 TTAGTTTATATTACTCTTATCTTTGAACATGAACATATGCCTTGTAGTGTTTATTGCT
 CAGCTGGCTGAGACACTGAAGAAGTCACTGAACAAAACCTACACACGTACCTTCACTGTGAT
 CACTGCTTCTCTCTTACCAGTCTATTCCACTGAACAAAACCTACACACATACCTTCAAT
 GTGGTTCAGTGCTTCTCTCTCTACCACTCTATTCCACTGAACAAAACCTACGACATAC
 CTTCATGTGGCTCAGTGCTTCTCTCTCTACCACTCTATTCCACTCTTTTCACTGTGTCT
 GACATGTTTGTGCTCTGTTCCATTAAACAACCTGCTCTTACTTCCAGTCTGTACAGAAATG
 CTATTTCACTTGAGCAAGATGATGTAATGAAAGGGTGTGGCACTGTGTCTGGAGACTGA
 TTTTGTAGTCTTGGTGCTATCAATCACCGTCTGTGTTGAGCAAGGCAATTTGGCTGCTGTA
 GCTTATTGCTTCACTGTGAAGCGGTGGTTTGAATTCCTGATCTTCCACCTCACAGTGATG
 TTGTGGGGTCCAGTGAGATAGAATACTGTAAGTGTGGTTTGAATTTAAAGGTGCTAT
 ACTAAGGGAAGAATTGAGGAATTAACGTCATACGTTTGGTGTGCTTTTTCAAGTGTGTA
 AAATAAAAAAATGTTAAG

10037391.102401

FIGURE 98

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52185

><subunit 1 of 1, 211 aa, 1 stop

><MW: 22744, pI: 8.51, NX(S/T): 1

MANAGLQLLGFI~~L~~AF~~L~~GWIGAI~~V~~STALPQWRIYSYAGDNIVTAQAMYEGLWMSCVSQSTGQI
QCKVFD~~S~~LLNLSSTLQATRALMVVGILLGVIAIFVATVGMKCMKCLEDDEVQKMRMAVIGGA
IFLLAGLAILVATAWYGNRIVQEFYDPMTPVNARYEFGQALFTGWAAAASLCLLGALLCCSC
PRKTTSYTPRPYPKPAPSSGKDYV

Important features:

Signal peptide:

amino acids 1-21

Transmembrane domains:

amino acids 82-102, 118-142 and 161-187

N-glycosylation site.

amino acids 72-75

PMP-22 / EMP / MP20 family proteins

amino acids 70-111

ABC-2 type transport system integral membrane protein

amino acids 119-133

FIGURE 99

TTCTGGCCAAACCCGGGGCTNCAGCTGTTGGGCTTCATCTCGCCTTCCTGGGATGGATCGGC
GCCATCNTCACACTGCCCTTCCCCAGTGGAGGATTTTACTCCCTATGCTGGCGACAACATCG
TGACCGCCCAGCCCATGTACGAGGGGCTGTGGATGTCNGCGTGTCGCAGAGCACCGGGCAG
ATCCAGTGCAAAGTCTTTGACTCCTTGCTGAATCTGAGCAGCACATTGCAAGCAAACCCGTGC
CTTGATGGTGGTTGGCATCCTCCTGGGAGTGATAGCAATCTTTGTGGCCACCGTTGGCATGA
AGTGATGAAGTGCTTGGAAGACGATGAGGTGCAGAAGATGAGGATGGCTGTCATTGGGGGC
GCGATATTTCTTCTTGCAAGTCTGGCTATTTTAGTTGCCACAGCATGGTATGGCAATAGAAN
CNTTCAACANTTCTATGACCCTATGACCCAGTCAATGCCAGGTACGAATTTGGTCA
GGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTGCTTCTGGGAGGTGCCCTACTTTGCT
GTTCTGTGCC

100
90
80
70
60
50
40
30
20
10
0

FIGURE 100

ACCCTTGACCCAACGCGGGCCCCCGACCGNTTCATGGCCAAACGCGGGNCTCCAGCTGTTGG
GCTTCATTCTCCCCCTTCCTGGGATGGACCGGCGCCCATCNTCAGCACTGCCCTGCCCCAGTG
GAGGATTTACTCCTATNCCGGCNACAACATCGTGACCGCCAGGCCNTGTACGAGGGGCTGT
GGATGTCCTGCGTGTGCGAGAGCACCGGGCAGATCCAGTGCAAAGTCTTTGACTCCCTTGCT
GAATCTGAGCAGCACATTGCAAGCAACCCGTGCCCTTGATGGTGGTTGGCATCCTCCTGGGAG
TGATAGCAATCTTNNNTGGCCACCGTTGTNNNTGAAGTGTATGAAGTGCTTGGAAGACGATGA
GGTGCAGAAGATGAGGATGGCTGTCATTGGGGGCGCGATATTTCTTCTTGCAAGTCTGGCTA
TTTTAGTTGCCACAGCATGGTATGGCAATAGAATCGTTCAAGAATTCTATGACCTATGACCGA

Figure 1 illustrates the stages of chick development from fertilization to hatching. The stages are numbered 1 through 12. Stage 1 shows a fertilized egg. Stage 2 shows cleavage. Stage 3 shows the formation of the blastoderm. Stage 4 shows gastrulation. Stage 5 shows the formation of somites. Stage 6 shows head development. Stage 7 shows heart development. Stage 8 shows lung development. Stage 9 shows liver development. Stage 10 shows kidney development. Stage 11 shows a chick in the egg. Stage 12 shows a hatched chick.

GGGCCCCGACCATTATCCAACCGGGNCTACTGTTGGCTCATCTCCCTCCTGGATGAANC GCGC
CATCNTCAGACTCCCTGCCCATGGAGATTNNCTATGCTGGCGACAACATCNTGACCCCC
AGCCATGTACGAGGGGCTTTGAACGTCTCGCGTGTGCGAGANACCGGGCAGATCCAGTGCAA
AGTCTTTGACTCCTTGCTGAATCTGNGCAGCACATTGCAGCAACCCNTGCGCTGATGGTGGT
TGGCATCTCCTGGGAGTGATAGCAATCTTTGTGGCCACCGTTGGCATGAAGTGATGAAGT
GCTTGGAAGACGATGAGGTGCAGAAGATGAGGATGGCTGTGTCATTGGGGGCGCGATATTTCTT
CTTGACAGGCTGTGGCTATTNNNGTTGCCACAGCATAGTGATGCGATAGAATCGTTCAAGAAT
TCTATGACCCCTATGACCCCACTCAATGCGCAGGTACGAAATTTGGTCAGGCTCTTTCACCTGGC
TGGGCTGCTGCTTCTCTCTGCTCTGGGAGGTGCCCTACTTTGCTGTTCTCTGCGA

FIGURE 102

ATTCTCCCCTCCTGGATGGATCGCNCCACCGTCACATTGCCTTCCCCANTGGAGGATTNAC
TCCTATGCTGGCGACAACATCGTGACCCCCAGGCCATTTACCGAGGGGCTTTGGATGTCNT
GCNTGTCGCAGAGCACCGGGCAGATCCCAGTGCAAAGTCTTTGACTCCTTGCTGAATCTGAG
CAGCACATTGCAAGCAACCCGTGCCTTGATGGGGTTGGCATCCTCCTGGGAGTGATAGCAAC
CTTTGTGGCCACCGTTGGCATGAAGTGTATGAAGTGCTTGAAGACGATGAGGTGCCAGAAG
ATGAGGATGGCTGTCAATTGGGGGCGCGATATTTCTTGTTGCAGGTCTGGCTATTTTAGTNGC
CACAGCATGGTATGGCAATAGANTNNTTCNNGNNNTCTATGACCCCTATGACCCAGTCAATG
CCAGGTACGAATTTGGTCAGGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTCTGCCTTCTG
GGAGGTGCCCTACTTTGCTGTTCTCTGTCCC

100
90
80
70
60
50
40
30
20
10
0

FIGURE 103

AGAGCACCGGCAGATCCCAGTNCAAAGTCTTTGACCCTTGCTGAATCTGAGCAGCACATTNC
AAGCAACCCCTTGCCCTTGAAGGTGGTTGNCATCCCCCTGGGAGTGAATAGCAATCTTTGTG
GCCACCGTTGGCATGAAGTNTATGAAGTGCTTGAAGACGATGAGGTGCAGAAGATGAGGAT
GGCTGTCAATTGGGGGCGCGATATTTCTTCTTGCAAGTCTGGCTATTTTAGTNNCCACAGCAT
GGTATGGCAATAGNATNNTTGNGGNTTCTATGACCCTATGACCCAGTCAATGCCAGGTAC
GAATTTGGTCAGGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTGCTTCTGCGCTTCTGGGAGGTGC
CCTACTTTGCTGTTCTCTGTCCCCGAA

FIGURE 104

AGCAATGCCCTGCCCCAGTGGAGGATTAAATTCCTATGNTGGGGACAACATTGTGACNGCCC
AGGCCATGTACGGGGGGCTGTGGATGTCTTGCCTGTGCGAGAGCACCGGGCAGATCCAGTGC
AAAGTNITTTGACTCCTTGCTGAATTTGAGCAGCACATTGCAAGCAACCCGTGCCTTGATGGT
GGTTGGCATCTTCCTGGGAGTGATAGCAATCTTTGTGGCCACCGTGGNAATGAAGTGTATGA
AGTGCTTGAAGACGATGAGGTGCAGAAGATGAGGATGGCTGTTCATTGGGGGCGCGATATTT
CTTNTTGCAGGTCTGGCTATTTTAGTTGCCACAGCATGGTATGGCAATAGAATNGTTCAAGA
ATTTTATGACCCTATGACCCAGTCAATGCCAGGTACGAATTTGGTCAGGCTTTNTTCACTG
GCTGGGCTGCTGCTTNTTCTGCCTTNTGGGAGGTGCCCTANTTTGCTGTTCTGCGAACC

$\left\{ \begin{array}{l} \mathbf{f}_1, \mathbf{f}_2 \\ \mathbf{f}_3, \mathbf{f}_4 \\ \mathbf{f}_5, \mathbf{f}_6 \\ \mathbf{f}_7, \mathbf{f}_8 \\ \mathbf{f}_9, \mathbf{f}_{10} \\ \mathbf{f}_{11}, \mathbf{f}_{12} \\ \mathbf{f}_{13}, \mathbf{f}_{14} \\ \mathbf{f}_{15}, \mathbf{f}_{16} \\ \mathbf{f}_{17}, \mathbf{f}_{18} \\ \mathbf{f}_{19}, \mathbf{f}_{20} \\ \mathbf{f}_{21}, \mathbf{f}_{22} \\ \mathbf{f}_{23}, \mathbf{f}_{24} \\ \mathbf{f}_{25}, \mathbf{f}_{26} \\ \mathbf{f}_{27}, \mathbf{f}_{28} \\ \mathbf{f}_{29}, \mathbf{f}_{30} \\ \mathbf{f}_{31}, \mathbf{f}_{32} \\ \mathbf{f}_{33}, \mathbf{f}_{34} \\ \mathbf{f}_{35}, \mathbf{f}_{36} \\ \mathbf{f}_{37}, \mathbf{f}_{38} \\ \mathbf{f}_{39}, \mathbf{f}_{40} \\ \mathbf{f}_{41}, \mathbf{f}_{42} \\ \mathbf{f}_{43}, \mathbf{f}_{44} \\ \mathbf{f}_{45}, \mathbf{f}_{46} \\ \mathbf{f}_{47}, \mathbf{f}_{48} \\ \mathbf{f}_{49}, \mathbf{f}_{50} \\ \mathbf{f}_{51}, \mathbf{f}_{52} \\ \mathbf{f}_{53}, \mathbf{f}_{54} \\ \mathbf{f}_{55}, \mathbf{f}_{56} \\ \mathbf{f}_{57}, \mathbf{f}_{58} \\ \mathbf{f}_{59}, \mathbf{f}_{60} \\ \mathbf{f}_{61}, \mathbf{f}_{62} \\ \mathbf{f}_{63}, \mathbf{f}_{64} \\ \mathbf{f}_{65}, \mathbf{f}_{66} \\ \mathbf{f}_{67}, \mathbf{f}_{68} \\ \mathbf{f}_{69}, \mathbf{f}_{70} \\ \mathbf{f}_{71}, \mathbf{f}_{72} \\ \mathbf{f}_{73}, \mathbf{f}_{74} \\ \mathbf{f}_{75}, \mathbf{f}_{76} \\ \mathbf{f}_{77}, \mathbf{f}_{78} \\ \mathbf{f}_{79}, \mathbf{f}_{80} \\ \mathbf{f}_{81}, \mathbf{f}_{82} \\ \mathbf{f}_{83}, \mathbf{f}_{84} \\ \mathbf{f}_{85}, \mathbf{f}_{86} \\ \mathbf{f}_{87}, \mathbf{f}_{88} \\ \mathbf{f}_{89}, \mathbf{f}_{90} \\ \mathbf{f}_{91}, \mathbf{f}_{92} \\ \mathbf{f}_{93}, \mathbf{f}_{94} \\ \mathbf{f}_{95}, \mathbf{f}_{96} \\ \mathbf{f}_{97}, \mathbf{f}_{98} \\ \mathbf{f}_{99}, \mathbf{f}_{100} \end{array} \right\}$

FIGURE 105

TCATAGGGGGGCGCGATATTTTTCTTGCAGGTTNTGGTTATTTTAGTTGCCACAGCATGGTA
TGGCAATAGAATCGTTCAAGAATTNTATGACCCTATGACCCCAGTCAATGCCAGGTACGAAT
TTGGTCAGGCTCTNTTCACTGGNTGGGCTGCTGCTTCTNTNNGCCTTNTGGGAGGTGCCCTA
CTTTGCTGTTCCCTG

100
90
80
70
60
50
40
30
20
10
0
100
90
80
70
60
50
40
30
20
10
0

FIGURE 106

TTCTCTGGGATGGATCCGCCCCCATCNTCACATGCCCTGCCCCNTGGAGATTACNCCTATGC
TGGCGAACACATCNTGACCGCCCAGGCCATGTACGAGGGGCTGTGGAATGTCTGCGTGTC
CCAGAGCACCGGGCAGATCCAGTGCAAAGTCTTTGACTCCTTGCTGAATCTGAGCAGCACAT
TGCAAGCAACCNTGCCCTTGATGGTGGTTGGCATCCTCCTGGGAGTGATAGCAATCTTTGTGG
CCACCGTTGGCATGAAAGTGATGAAAGTGCTTGGAAGACGATGAGGTGCAGAAGATGAGGAT
GGCTGTCAATTGGGGGCGCGATATTTCTTCTTGCAAGTCTGGCTATTTTAGNNGCCACAGCAT
GGTATGGCAATCAGACCCNNTCANAACTCTATGACCCCTATGACCCAGTCAATGCCAGGTA
CGAATTTGGTCAGGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTCTGCCCTTCTGGGAGGTG
CCCTACTTTGCTGTTCTGTCCCCGAAAAACAACCTCTTACCCACG

100
90
80
70
60
50
40
30
20
10
0

FIGURE 107

CGGGGCTGCAGCTGTTGGGCTTCATCTCGCTTCCTGGGATGGAATCGGCGCCATCGTCAGCA
CTGCCCCTGCCCCATGGAGGATTTACTCNTATGCTGGCGACAACATCGTGACCNCCCAGGCCA
TGTACGAGGGGCTGTGGATGTCNGCGTGTGCGAGAGCACCGGGCAGATCCAGTGCAAAGTCT
TTGACTCCTTGCTGAATCTGAGCAGCACATTGCAAGCAACCNTGCCTTGATGGTGGTTGGCA
TCCTCCTGGGAGTGATAGCAATCTTTGTGGCCACCGTTGGCATGAAGTGATGAAGTGCTTG
GAAGACGATGAGGTGCAGAAGATGAGGATGGCTGTCATTGGGGGCGCGATATTTCTTCTTGC
AGGTCTGGCTATTTNTAGTTGCCACAGCATGGTATGGCAATAGAATCGTTCAAGAATTCTAT
GACCCTATGACCCAGTCAATGCCAGGTACGAATTTGGTCAGGCTCTCTTCACTGGCTGGGC
TGCTGCTTCTCTCTGCCTTCTGGGAGGTGCCCTACTTTGCTGTTCTCTGCGAA

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

FIGURE 108

GCCTGCCGTAGCTCGCCGGGACCGCGGCTCGCCCTCGCCCTCCGCCCCGCGCCTGCAC
CGCGTAGACCGACCCCCCTCCAGCGCGCCACCGGTAGAGGACCCCCCGCGCTGCCCG
ACCGGTCCCCGCGCTTTTGTAAAACTTAAAGCGGGCGCAGCATTAAACGCTTCCCGCCCCGT
GACCTCTCAGGGGTCTCCCGCCAAAGGTGCTCCGCGCTAAGGAACATGGCGAAGGTGGAG
CAGGTCCTGAGCCTCGAGCGCAGCAGAGCTCAAATTCGAGGTCCTTCCCGATGTTGT
CACCACCAACCTTAAAGCTTGGCAACCCGACAGACCGAAATGTGTGTTTAAAGTGAAGACTA
CAGCACCACTAGGTACTGTGTGAGGCCCAACAGCGGAATCATCGATGCAGGGGCTCAATT
AATGTATCTGTGATGTACAGCCTTTCGATTATGATCCCAATGAGAAAAGTAAACACAAGTT
TATGGTTCAGTCTATGTTTGTCTCAACTGACACTTCAGATATGGAAGCAGTATGGAAGGAGG
CAAAACCGGAAGACCTTATGGATTCAAACTTAGATGTGTGTTTGAATTGCCAGCAGAGAAT
GATAAACCATGATGATAGAAATAAATAAAATATATCCCAACTGCATCAAAGACAGAAAC
ACCAATAGTGTCTAAGTCTCTGAGTTCCTCTTGGATGACACCGAAGTTAAGAAGGTTATGG
AAGAATGTAAGAGGCTGCAAGGTGAAGTTCAGAGGCTACGGGAGGAGAACAGCAGTTCAG
GAAGAAGATGGACTGCGGATGAGGAAGACAGTGCAGAGCAACAGCCCCATTTCAGCATTAGC
CCCAACTGGGAAGGAAGAAGGCTTAGCACCCGGCTCTTGGCTCTGGTGGTTTGTCTTTA
TCGTTGGTGTAATTATTGGGAAGATTGCCTTGTAGAGGTAGCATGCACAGGATGGTAAATTG
GATTGGTGGATCCACCATATCATGGGATTAAATTTATCATAACCATGTGTAAAAAGAAATT
AATGTATGATGACATCTCACAGGTCTTGCCCTTAAATTACCCCTCCCTGCACACATACAC
AGATACACACACAAATATAATGTAACGATCTTTAGAAAAGTTAAAAATGTATAGTAACTG
ATTGAGGGGAAAAAGAATGATCTTTATTAATGACAAAGGAAACCATGAGTAATGCCACAAT
GGCATATTGTAAATGTCATTTTAAACATTGGTAGGCCCTTGGTACATGATGCTGGATTACCTC
TCTTAAATGACACCCCTTCCTCGCTGTGGTGTCTGGCCCTTGGGGAGCTGGAGCCGAGCAT
GCTGGGGAGTGGGTGAGCTCCACACAGTAGTCCCAAGTGGCCCACTCCCGGCCAGGCTG
CTTCCGTCTCTTCAGTTCGTGCCAAGCCATCAGCTCCTTGGGACTGATGAACAGAGTTCAGA
AGCCCAAAGGAATTGCACTGTGGCAGCATCAGACGTACTCGTCATAAGTCAGAGGCGTGTGT
TGACTGATTGACCCAGCGCTTGGAAATAAATGGCAGTGTCTTGTTCACCTAAAGGGACCAA
GCTAAATTTGATTTGGTTCATGTAGTGAAGTCAAACCTGTTATTCAGAGATGTTTAAATGCATA
TTTAACTTATTTAATGTATTTTATCTCATGTTTTCTTATGTGACAAGAGTACAGTTAATGC
TGCGTGCTGCTGAACCTGTTGGGTGAACCTGGTATTGCTGCTGGAGGGCTGTGGGCTCCTCT
GTCTCTGGAGAGTCTGGTTCATGTGGAGGTGGGGTTTATGGGATGCTGGAGAAGAGCTGCCA
GGAAGTGTTTTTCTGGGTGAGTAAATAACAACCTGTATAGGGAGGGAATTCAGTAGTG
ACAGTCAACTCTAGGTTACCTTTTTTAAATGAAGAGTAGTCAGTCTTCAGATTGTTCTTATA
CCACCTCTCAACCATTACTCACACTCCAGCGCCAGGTCCAAGTCTGAGCCTGACCTCCCC
TTGGGACCTAGCCTGGAGTCAGGACAAATGGATCGGGCTGCAGAGGGTTAGAAGCGAGGGC
ACCAGCAGTTGTGGGTGGGGAGCAAGGGAAGAGAGAACTCTTCAGCGAATCCTTCTAGTAC
TAGTTGAGAGTTTGAATTAATTTTATGCCATAAAAGACCAACCCAGTTCTGTGTTGA
CTATGTAGCATCTTGAAGAAAAAATTATAATAAAGCCCCAAATTAAGAAA

10017991.102403

FIGURE 109

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53977

<subunit 1 of 1, 243 aa, 1 stop

<MW: 27228, pI: 7.43, NX(S/T): 2

MAKVEQVLSLEPQHELKFRGPFPTDVVTNLKLGNPTRDNVCFKVKTTPRRYCVRPNSGIID
AGASINVSVMQLQPFDDPNEKSKHKFMVQSMFAPTTDSMEAVWKEAKPEDLMDSKLRVFE
LPAENDKPHDVEINKIISTASKTETPIVSKSLSSSLDDTEVKKVMEECKRLQGEVQRLREE
NKQFKEEDGLRMRKTVQSNSPISALAPTGKEEGLSTRLLALVVLFFIVGVIIGKIAL

Important features:

Transmembrane domain:

amino acids 224-239

N-glycosylation site.

amino acids 68-71

N-myristoylation site.

amino acids 59-64, 64-69 and 235-240

FIGURE 110

GTCAGTCTTCTAGATTGTCCTTATCCACCTTTCAACCANTACTCACATTTTCNAGCGCCCAG
GTCCANGTCTGAGCCTGACTTCCCCTTGGGGACCTAGCCTGGAGTCAGGACAATGGNTCGGG
CTGCAGAGGNTTAGAAGCGAGGGCACCAGCAGTTTTGGGTGGGGAGCAAGGGNNGAGAGAAA
CTCTTCAGCGAATCCTTCTAGTACTAGTTGAGAGTTTGACTGTGAATTAATTTTATGCCATA
AAAGACNAACCCAGTTCTGTTTGA CTATGTAGCATCTTGAAAAGAAAAATTATAATAAGCC
CCAAAATTAAGAATCTTTTGTCA'TTTTGTACATTTGCTCTATGGGGGAATTATTATTTT
ATCATT'TTTATTATTTTGCCATTGGAAGGTAACTTTAAAATGAGC

TATTGTAAAGGCCATTTTAAACCATTGGTAGGCCTGGTACATGATGCTGGATTACCTCCTT
AAATGACACCNTTCTCGCCTGTTGGTGCTGGCCNTTGGGGAGCTGGAGCCCCAGCATGCTG
GGGAGTGCGGTCAGCTCCACACAGTAGTCCCCAGTGGCCCACTCCCGGCCAGGCTGCTTT
CCGTGTC'TTCAGTTCTGTCCAGCCATCAGCTCCTTGGGACTGATGAACAGAGTCAGAAGCC
CAAAGGAATTGCCACTGTGGCAGCATCAGACGTACTCGTCATAAGTGAGAGGCGTGTGTTGA
CTGATTGACCCAGCGCTTTGGAAATAAATGGCAGTGCTTTGTTCACTTAAAGGGACCAAGCT
AAATTGATTTGGTTTCATGTAGTGAAGTCAAAGTCTTATTCAGAGATGTTTAAATGCATATTTA
ACTTATTTAATGTATTTTCATCTCATGTTTTCTTATTGTCACAAGAGTACAGTTAATGCTGCG
TGCTGCTGAACTCTGTTGGGTGAACTGGTATTGCTGCTGGAGGGCTG

FIGURE 112

CCCTGGTGGTTTTGTTCTTTAATTCGTTGGTGTAATTNTTGGGAAGATTGCTTGTAGAGGTA
GNATGCACCNCGGCTGGTAAATTGGATTGGTGGATCCACCATATCCATGGGATTTAAATTTAT
CATAACCATGTGTAAAAAGAAATTAATGTATGATGACATNTCACAGGTATTGCCTTTAAATT
ACCCATCCCTGNANACACATACACAGATACACANANACAAATNTAATGTAACGATNTTTTAG
AAAGTTAAAAATGTATAGTAAC

CCCTGGTGGTTTTGTTCTTTAATTCGTTGGTGTAATTNTTGGGAAGATTGCTTGTAGAGGTA

FIGURE 113

GGTGGCCCATTCCTGGGCCAGGCTGCTTTCGGTNTTCAGTTCTGTCCAAGCCATCAGCTCC
TTGGGACTGATGAACAGAGTCAGAAGCCCAAAGGAATTGCACTGTGGCAGCATNAGACGTAC
TTGTNATAAGTGAGAGGCGTGTGTGACTGATTGACCCAGCGCTTTGAAAATAAATGGCAGT
GCTTTGTTTCANTTAAAGGGACCAAGCTAAATTGTATTGGTTCATGTAGTGAAGTCAAACCTG
TTATTCAGAGATGTTTAAATGCATATTTAANTTATTAAATGTATTNATNTCATGTTTTCTTA
TTGTCACAAGAGTACAGTTAATGCTGCGTGCTGCTGAANTNTGTTGGGTGAACCTGGTATTGC
TGCTGGAGGGCTGTGGGCTCCTCTGTCTTTGGAGAGTCTGGTCATGTGGAGGTGGG

FIGURE 114

TGCTTCCGTGCTTCAGTTCGTCCAAGCCATCAGCTCCTTGGGACTTGATGAACAGAGTC
AGAAGCCCAAAGGAATTGCACTGTGGCAGCATCAGACGTACTCGTCATAAGTGAGAGGCGTG
TGTTGACTGATTGACCCAGCGCTTGGAAATAAATGGCAGTGCTTTGTTCACTTAAAGGGAC
CAAGCTAAATTTGTATTGGTTCATGTAGTGAAGTCAAACGTATTATTGAGAGATGTTTAATGC
ATATTTAACTTATTTAATGTATTTTCATCTCATGTTTCTTATTGTCACAAGAGTACAGTTAA
TGCTGCGTGC

10037191-202401

FIGURE 115

AAACCTTTAAAAGTTGAGGGGAAAAGAATGATCCTTTATTATGACAAGGGAAACNTGNGT
AATGCCACAATGGCATATTGTAATGTCATTTTAAACATTGGTAGGCCTTGGTACATGATGC
TGGATTACCTCTCTTAAAATGACACCCCTTCCTCGCCTGTTGGTGCTGGCCCTTGGGGAGCTN
GAGCCCAGCATGCTGGGGAGTGCGGTCTGCTCCACACAGTAGTCCCCANGTGGCCANTCCC
GGCCCAGGCTGCTTTCCGTGCTTTCAGTTCTGTCCAAGCCATCAGCTCCTTGGGANTGATGA
ACAGAGTCAGAAGCCCAAAGGAATTGCANTGTGGCAGCATCAGANGTANTNGTCATAAGTGA
GAGGCGTGTGTTGANTGATTGACCCAGCGCTTTGGAAATAAATGGCAGTGCTTTGTTCANTT
AAAGGNC CAAGNTAAATTTGTATTGGTTCATGTAGTGAAGTCAAANTGTTATTTCAGAGATG
TTTAATGCATATTTAANTTATTTAATGTATTTCATNTCATGTTTTCTTATTGTCACAAGGGT
ACAGTTAATGCTGCGTGCTGCTGAANTCTGTTGGGTGAANTGGTATTGCTG

FIGURE 116

GGCCCTTGGGGAGCTGGAGCCCAGCATGCTGGGGAGTGCGGTCAGCTCCACACAGTAGTCCC
CACGTGGCCCACTCCCGGCCAGGCTGCTTTCCGTGTCTTCAGTTCGTGCCAAGCCATCAGC
TCCTTGGGACTGATGAACAGAGTCAGAAGCCCAAAGGAATTGCACTGTGGCAGCATCAGACG
TACTCGTCATAAGTGAGAGGCGTGTGTTGACTGATTGACCCAGCGCTTTGGAAATAAATGGC
AGTGCTTTGTTCACCTTAAGGGACCAAGCTAAATTTGTATTGGTTCATGTAGTGAAGTCAAA
CTGTTATTCAGAGATGTTTAATGCATATTTAACTTATTTAATGTATTTTCATCTCATGTTTTC
TTATTGTCACAAGAGTACAGTTAATGCTGCGTGCTGCTGAACTCTGTTGGGTGAACTGGTAT
TGCTGCTGGAGGGCTGTGGGCTCCTCTGTCTCTGGAGAGTCTGGTCATGTGGAGGTGGG

FIGURE 117

GCGAGCTCCGGGTGCTGTGGCCCGGCCTTGGCGGGGCGGCCTCCGGCTCAGGCTGGCTGAGA
 GGCTCCCAGCTGCAGCGTCCCCGCCCGCCTCCTCGGGAGCTCTGATCTCAGCTGACAGTGCC
 CTCGGGGACCAAAACAAGCCTGGCAGGGTCTCACTTTGTTGCCAGGCTGGAGTTCAGTGCCA
 TGATCATGGTTTACTGCAGCCTTGACCTCCTGGTTCAAGCGATCCTGCTGAGTAGCTGGGA
 CTACAGGACAAAATTAGAAGATCAAAATGGAAAATATGCTGCTTTGGTTGATATTTTACC
 CCTGGGTGGACCTCATTGATGGATCTGAAATGGAATGGGATTTTATGTGGCACTTGAGAAA
 GGTACCCCGGATTGTCAGTGAAAGGACTTTCCATCTCACCAGCCCGCATTTGAGGCAGATG
 CTAAGATGATGGTAAATACAGTGTGTGGCATCGAATGCCAGAAAGAACTCCCAACTCCCAGC
 CTTTCTGAATTGGAGGATTATCTTTCCCTATGAGACTGTCTTTGAGAATGGACCCGAACCTT
 AACCGGGTGAAAGTTCAAGATTGTGTTCTTGAGCCGACTCAAAATATCACCACAAGGGAG
 TATCTGTTAGGAGAAAGAGACAGGTGTATGGCACCGACAGCAGGTTCAGCATCTTGGACAAA
 AGGTTCTTAAACCAATTTCCCTTTCAGCACAGCTGTGAAGCTTTCCACGGGCTGTAGTGGCAT
 TCTCATTTCCTCAGCATGTTCTAACTGCTGCCACTGTGTTATGATGGAAGGACTATG
 TCAAAGGGGATAAAAGCTAAGGGTAGGGTTGTTGAAGATGAGGAATAAAGTGGAGGCAAG
 AAACGTCGAGGTTCTAAGAGGAGCAGGAGAGAAGCTAGTGGTGGTGACCAGAGAGGGGTAC
 CAGAGAGCATCTGCAGGAGAGAGCGAAGGGTGGGAGAGAAGAAAAAATCTGGCCGGGGT
 AGAGGATTGCCGAAGGGAGGCGCTTCTTTCAGTGGACCCGGTCAAGAATACCCACATTCCG
 AAGGGCTGGGCACGAGGAGGATGGGGGACGTACCTTGGACTATGATATGCTCTTCTGGA
 GCTGAAGCGTGCTCACAAAAAGAAATACATGGAACCTTGAATCAGCCCAACGATCAAGAAAA
 TGCCTGGTGAATGATCCACTTCTCAGGATTGATAACGATAGGGCTGATCAGTTGGTCTAT
 CGGTTTTCAGTGCTGTCGAGCAATCCAATGATCTCCTTTACCAATACTGCGATGCTGAGTC
 GGGCTCCACCGGTTCCGGGGTCTATCTGCGTCTGAAAGATCCAGACAAAAGAATTGGAAGC
 GCAAAATCATTTGCGGTCTACTCAGGGCACCAAGTGGGTGGATGTCCACGGGGTTCAAGAGGAC
 TACAACGTTGCTGTTGCGATCACTCCCTAAAATACGCCAGATTGCTCTGGATTCAAGG
 GAACGATGCCAATTTGTGCTTACGGCTAAACAGAGACCTGAAACAGGGCGGTGTATCATCTAAA
 TCACAGAGAAAAACAGCTCTGCTTACCGTAGTGAGATCACTTATAGGTTATGCTGGACTT
 GAACTCTGCAATAGCATTTCAACATTTTTCAAATCAGGAGATTTTCGTCCATTTAAAAAA
 TGTATAGGTGCAGATATTGAAACTAGGTGGGCATTTCAATGCCAAGTATATACTCTTCTTTA
 CATGGTGATGAGTTTCATTGTGAGAAAAATTTTGTGTCCTTCTTAAAAATTAGACACACTTT
 AAACCTTCAAACAGGTATTATAAATAACATGTGACTCCTTAATGGACTTATTCTCAGGGTCC
 TACTCTAAGAAGAACTAATAGGATGCTGGTTGTGATTAAATGTGAAATTGCATAGATAAA
 GGTAGATGGTAAAGCAATTAGTATCAGAATAGAGACAGAAAGTTACAACACAGTTTGTACTA
 CTCTGAGATGGATCCATTCACTCATGCCCTCAATGTTTATATTGTGTTATCTTGGGTCT
 GGGACATTTAGTTTAGTTTTTTTGAAGAAATACAAATCAGAAGAAAAAGCAAGCATTATAAA
 CAAAACATAAAGTCTTTTACTGCTTTAAGAAATAACAAATTACAATGTGTTATTTAAAAA
 TGGAGAAATAGTTTTGTTCTATGAAATAAACCTAGTTTGAAGAAATAGGAAAGCAGCATTT
 TAAGATCTCAAGTTTTTATTTAACTAATACTCAAAATATGGACTTTTCAATGATGCATAGGG
 AAGACACTTCACAAATTAGATGATCATGTGTTGAAAGCCACATTTTTATGCTATACAT
 TCTATGTATGAGGTCTACATTTTTTAGGACAAAGAAATCTGTAATCTTTTTCAAGAAAGAGT
 CTTTTCTCCTTGACAAATCCAGCTTTGTATGAGGACTATAGGGTGAATCTCTGATTAG
 TAATTTTAGATATGCTTTCCTAAAAATGAATAAAATTTATGAATATGA

FIGURE 118

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57253

<subunit 1 of 1, 413 aa, 1 stop

<MW: 47070, pI: 9.92, NX(S/T): 3

MENMLLWLIFFTPGWTLIDGSEMEWDFMWHLRKVPRIVSERTFHLTSPAFEADAKMMVNTVC
GIECQKELPTPSLSELEDYLSYETVFENGTRTLTRVKVQDLVLEPTQNITTKGVSVRRKRQV
YGTDSRFSIILDKRFLTNFPFSTAVKLSTGCSGILISQHVLTAAHCVDHGKDYVKGSKKLRV
GLLKMRNKS GGGKKRRGSKRSRREASGGDQREGTREHLQERAKGGRRRKSGRGQRIAEGRPS
FQWTRVKNTHIPKGWARGGMG DATLDYDYALLELKRAHKKKYMELGISPTIKKMPGGMIHFS
GFDNDRADQLVYRFCVSVSDESNDLLYQYCDAESGSTGSGVYLRLKDPDKKNWKRKIIAVYSG
HQWVDVHGVQKDYNNVAVRITPLKYAQICLWIHGNDANCAYG

Important features:

Signal peptide:

amino acids 1-16

N-glycosylation sites.

amino acids 90-93, 110-113 and 193-196

Glycosaminoglycan attachment site.

amino acids 236-239

Serine proteases, trypsin family, histidine active site.

amino acids 165-170

FIGURE 119

AATGTGAGAGGGGCTGATGGAAGCTGATAGGCAGGACTGGAGTGTAGCACCAAGTACTGGAT
GTGACAGCAGGCAGAGGAGCACTTAGCAGCTTATTCAGTGTCCGATTCTGATTCCGGCAAGG
ATCCAAGCATGGAATGCTGCCCTCGGGCAACTCCTGGCAGCTGCTCCTCTTTCTGGCTTTC
CTGCTCCTGAGTTCCAGGACCGCACGCTCCGAGGAGGACCGGGACGGCCTATGGGATGCCTG
GGGCCCCATGGAGTGAATGCTACGCACCTGCGGGGAGGGGCCCTCTACTCTCTGAGGCGCT
GCCTGAGCAGCAAGAGCTGTGAAGGAAGAAATATCCGATACAGAACATGCAGTAATGTGGAC
TGCCCAACCAGAAGCAGGTGATTTCCGAGCTCAGCAATGCTCAGCTCATAATGATGTCAAGCA
CCATGGCCAGTTTTATGAATGGCTTCCTGTGTCTAATGACCCTGACAACCCATGTTCACTCA
AGTGCCAAGCCAAAGGAACAACCCCTGGTTGTTGAACTAGCACCTAAGGTCTTAGATGGTACG
CGTTGCTATACAGAATCTTTGGATATGTGCATCAGTGGTTTATGCCAAATTTGTTGGCTGCGA
TCACCAGCTGGGAAGCACCCGTCAAGGAAGATAACTGTGGGGTCTGCAACGGAGATGGGTCCA
CCTGCCGGCTGGTCCGAGGGCAGTATAAATCCCAGCTCTCCGCAACCAATCGGATGATACT
GTGGTTGCCTTCCCTATGGAAGTAGACATATTCGCCTTGCTCTTAAAGGTCTGTATCACTT
ATATCTGGAACCAAAACCCCTCCAGGGGACTAAAGGTGAAAACAGTCTCAGCTCCACAGGAA
CTTTCCTTGTGGACAATTTCTAGTGTGGACTTCCAGAAATTTCCAGACAAAGAGATACTGAGA
ATGGCTGGACCACTCAGCAGAGATTTTCTATCAACCCATCATCCACCGATGGAGGAGACGGATTTCTTTC
CTTGCTCAGCAACCTGTGGAGGAGGTTATCAGCTGACATCGGCTGAGTGTCTACGATCTGAGG
AGCAACCGTGTGGTTGCTGACCAATACTGTCACTATTACCCAGAGAACATCAAAACCAAAACC
CAAGCTTCAGGAGTGCAACTTGGATCCTTGTCCAGCCAGTGACGGATACAAGCAGATCATGC
CTTATGACCTCTACCATCCCCTTCTCGGTGGGAGGCCACCCCATGGACCGCGTGTCTCTCC
TCGTGTGGGGGGGGCATCCAGAGCCGGGCAGTTTTCTGTGTGGAGGAGGACATCCAGGGGCA
TGTCACCTCAGTGAAGAGTGGAAATGCATGTACACCCCTAAGATGCCATCGCGCAGCCCT
GCAACATTTTTGACTGCCCTAAATGGCTGGCACAGGAGTGGTCTCCGTGCACAGTGACATGT
GGCCAGGGCCTCAGATACCGTGTGGTCTCTGCATCGACCATCGAGGAATGCACACAGGAGG
CTGTAGCCCAAAACAAAGCCCCACATAAAAGAGGAATGCATCGTACCCACTCCCTGTCTATA
AACCCAAAGAGAAATTTCCAGTCGAGGCCAAGTTGCCATGGTTCAAAACAAGCTCAAGAGCTA
GAAGAAGGAGCTGCTGTGTGTCAGAGGAGCCCTCGTAAAGTTGTAAGCAAGCTCAAGAGCTA
TTTGAACTGTTTTGTTTAAAGAAAGCAGTGTCTCACTGGTTGTAGCTTTTATGGGTTCTGA
ACTAAGTGTAATCATCTCACCAGGCTTTTTGGCTCTCAAAATAAGATTGATTAGTTTCAAA
AAAAAAAA

FIGURE 120

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58847

<subunit 1 of 1, 525 aa, 1 stop

<MW: 58416, pI: 6.62, NX(S/T): 1

MECCRRATPGTLLLFLLSSRTARSEEDRDGLWDAGWPWSECSTCGGGASYSLRRCLS
SKSCEGRNIRYRTCSNVDCPPEAGDFRAQQCSAHNDVKHHGQFYEWLPVSNPDNPNCSLKQC
AKGTTLVVELAPKVLDTGTRCYTESLDMCISGLCQIVGCDHQLGSTVKEDNCGVCNGDGSTCR
LVRGQYKSQLSATKSDDTVVALPYGSRHIRLVLKGPDHLYLETKTLQGTKGENSLSSTGTFL
VDNSSVDFQKFPDKEILRMAGPLTADFIVKIRNSGSADSTVQFIFYQPIIHRWRETDFFPSCS
ATCGGGYQLTSAECYDLRSNRVVADQYCHYYPENIKPKPKLQECNLDPCPASDGYKQIMPYD
LYHPLPRWEATPWTACSSSCGGGIQSRVSCVEEDIQGHVTSVEEWKCMYTPKMPIAQPCNI
FDCPKWLAQEWSPCTVTGCGGLRYRVVLCIDHRGMHTGGCSPKTKPHIKEECIVPTPCYKPK
EKLPEAKLPWFKQAQEELEGA AVSEEPS

Important features:

Signal peptide:

amino acids 1-25

N-glycosylation site.

amino acids 251-254

Thrombospondin 1

amino acids 385-399

von Willebrand factor type C domain proteins

amino acids 385-399, 445-459 and 42-56

FIGURE 121

CGGACGCGTGGGCGGCGGCTGCGGAACCTCCCGTGGAGGGGCCGGTGGGGCCCTCGGGCCTGAC
AGATGGCAGTGGCCACTGCGGCGGCAGTACTGGCCGCTCTGGGCGGGGCGCTGTGGCTGGCG
GCCCGCCGGTTCGTGGGGCCAGGGTCCAGCGGCTGCGCAGAGGCGGGGACCCCGGCCCTCAT
GCACGGGAAGACTGTGCTGATCACCGGGGCGAACAGCGGCCTGGGCGCGCCACGGCCGCGG
AGCTACTGCGCCTGGGAGCGCGGGTGATCATGGGCTGCCGGGACCGCGCGCGCCGAGGAG
GCGGCGGGTCAGCTCCGCGCGAGCTCCGCCAGGCCGCGGAGTGCGGCCAGAGCCTGGCGGT
CAGCGGGGTGGGCGAGCTCATAGTCCGGGAGCTGGACCTCGCCTCGCTGCGCTCGGTGCGCG
CCTTCTGCCAGGAAATGCTCCAGGAAGAGCCTAGGCTGGATGTCTTGATCAATAACGCAGGG
ATCTTCCAGTGCCCTTACATGAAGACTGAAGATGGGTTTGAGATGCAGTTCGGAGTGAACCA
TCTGGGGCACTTTCTACTACCAATCTTCTCCTTGGACTCCTCAAAAGTTCAGCTCCCAGCA
GGATTGTGGTAGTTTCTTCCAAACTTTATAAATACGGAGACATCAATTTTGATGACTTGAAC
AGTGAACAAGCTATAATAAAAGCTTTTGTATAGCCGGAGCAAACCTGGCTAACATTTCTTTT
TACCAGGGAAGTAGCCCGCCGCTTAGAAGGCACAAATGTCAACGTCAATGTGTGTCATCCTG
GTATTGTACGGACAAATCTGGGGAGGCACATACACATTCCACTGTTGGTCAAACCACTCTTC
AATTTGGTGTGATGGGCTTTTTTCAAACCTCCAGTAGAAGGTGCCCAGACTTCCATTTATTT
GGCCTCTTCACCTGAGGTAGAAGGAGTGTGAGGAAGATACCTTTGGGGATTGTAAGAGGAAG
AACTGTTGCCCAAAGCTATGGATGAATCTGTTGCAAGAAAACCTCTGGGATATCAGTGAAGTG
ATGGTTGGCCTGCTAAAAATAGGAACAAGGAGTAAAAGAGCTGTTTATAAACTGCATATCAG
TTATATCTGTGATCAGGAATGGTGTGGATTGAGAACTTGTTACTTGAAGAAAAAGAAATTTTG
ATATTGGAATAGCCTGCTAAGAGGTACATGTGGGTATTTTGGAGTTACTGAAAAATATTTTT
TGGGATAAGAGAATTTTCCAGCAAAGATGTTTTAAATATATATAGTAAGTATAATGAATAATAA
GTACAATGAAAAATACAATTATTTGTAATAATTATACTGGGCAAGCATGGATGACATATTA
ATATTGTGAGAAATTAAGTGACTCAAAGTGCTATCGAGAGGTTTTTCAAGTATCTTTGAGTT
TCATGGCCAAAGTGTTAACTAGTTTACTACAATGTTTGGTGTGTTGTGTGGAATATATCTGC
CTGGTGTGTGCACACAAGTCTTACTTGAATAAATTTACTGGTAC

FIGURE 122

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58747

<subunit 1 of 1, 336 aa, 1 stop

<MW: 36865, pI: 9.15, NX(S/T): 2

MAVATAAAVLAAALGGALWLAARRFVGPRVQRLRRGGDPGLMHGKTVLITGANSGLGRATAAE
LLRLGARVIMGCRDRARAEEAAGQLRRELRLQAACGPEPGVSGVGELIVRELDLASLRSVRA
FCQEMLQEEPRLDVLINNAGIFQCPYMKTEDGFEMQFGVNHLLGHFLLTNLLLGLLKSSAPSR
IVVSSKLYKYGDINFDDLNSEQSYNKSFCYSRSKLANILFTRELARRLEGNTVTVNLHPPG
IVRTNLGRHIHPLLVKPLFNLVSWAFFKTPVEGAQTSIYLASSPEVEGVSGRYFGDCKEEE
LLPKAMDESVARKLWDISEVMVGLLK

Important features:

Signal peptide:

amino acids 1-21

Short-chain alcohol dehydrogenase family protein

amino acids 134-144, 44-56 and 239-248

N-glycosylation site.

amino acids 212-215 and 239-242

Abstract

GGGGATTGTAAGAGGAGAGNACTGTGCCCAAAGNTATGGATGAATCTGTTGCAAGAAAATTN
TGGGATATCAGTGAAGTGATGTTNGCCTGCTAAAAATAGGAACAAGGAGTAAAAGAGCTGTT
TATAAAATGTCATATCAGTTATATCTGTGATCAGGAATGGTGTGGATTGAGAACCTTGTTACT
TGAAGAAAAAGAAATTTTGATATTGGAATAGCCTGNTAAGAGGNA CATGTGGGTATTTTGGAG
TTACTGAAAAATTATTTTGGGATAAGAGAATTCAGCAAAGATGTTTTAAATATATATAGT
AAGTATAATGAATAATAAGTACAATGAAAAATACAATTATATTTGTAATAATACTGGGCA
AGCATGGATGACATATTAATTAATTTGTCAGAATTAAGTGACTCAAAGTGCTATCGAGAGGTTT
TTCAGTATCTTCTAGTTTTCATGCGCCAAAGTGTTAACTAGTTTTACTACAATGTTTGGTGTT
TGTGTGAAATTATCTGCCTGGCTT

FIGURE 124

GAGAGGACGAGGTGCCGCTGCCTGGAGAATCCTCCGCTGCCGTCCCGGAGCCCAGCC
CTTTCCTAACCCAAACCAACCTAGCCCAGTCCCAGCCGCCAGCGCCTGTCCCTGTACAGGAC
CCCAGCGTTACCAATGCATCTCTGCCGTCTTCTATCCTTACC CGACCTCAGATGCTCCCTTCT
GCTCCTGGTAACTTGGGTTTTACTCCTGTAAACAACTGAAATAACAAGTCTTGCTACAGAGA
ATATAGATGAAATTTTAAACAATGCTGATGTTGCTTTAGTAAATTTTTATGCTGACTGGTGT
CGTTTCAGTCAGATGTTGCATCCAATTTTTGAGGAAGCTTCCGATGTCATTAAGGAAGAATT
TCCAAATGAAATCAAGTAGTGTGGCCAGAGTTGATTGTGATCAGCACTCTGACATAGCCC
AGAGATACAGGATAAGCAAATACCCAACCTCAAATTGTTTCGTAATGGGATGATGATGAAG
AGAGAATACAGGGGTCAGCGATCAGTGAAAGCATTGGCAGATTACATCAGGCAACAAAAAG
TGACCCCATTCAAGAAATTCGGGACTTAGCAGAAATCACCACCTCTTGATCGCAGCAAAAGAA
ATATCATTGGATATTTTGAGCAAAAGGACTCGGACAACATATAGAGTTTTTGAAACGAGTAGCG
AATATTTTGCATGATGACTGTGCCTTTCTTCTGCATTTGGGGATGTTCAAAACCGGAAAG
ATATAGTGGCGACAACATAATCTACAAACCACCAGGGCATTCTGCTCCGATATGGTGACT
TGGGAGCTATGACAAATTTTGATGTGACTTACAATTGGATTCAAGATAAATGTGTTCTCTT
GTCGAGAAATAACATTTGAAATGGAGAGGAATTGACAGAAGAAGACTGCCTTTTCTCAT
ACTCTTTCACATGAAAGAAGATACAGAAAGTTAGAAATATTCCAGAATGAAGTAGCTCGGC
AATTAATAAGTGAAAAGGTACAATAAACTTTTTACATGCCGATTGTGACAAATTTAGACAT
CCTCTTCTGCACATACAGAAAACCTCCAGCAGATTGTCTGTAAATCGCTATTGACAGCTTTAG
GCATATGTATGTGTTTGGAGACTTCAAAGATGTATTAATTCCTGGAAAACCAAGCAATTG
TATTTGACTTACATTCTGAAAACCTGCACAGAGAATTCATCATGGACCTGACCCAACTGAT
ACAGCCCCAGGAGAGCAAGCCCAAGATGTAGCAAGCAGTCCACCTGAGAGCTCCTTCCAGAA
ACTAGCACCCAGTGAATATAGGTATACTCTATTGAGGGATCGAGATGAGCTTTAAAAAAAACTTG
AAAAACAGTTTGTAAGCCTTTCAACAGCAGCATCAACCTACGTGGTGAAATAGTAAACCTA
TATTTTCATAATTCTATGTGTATTTTTATTTTGAATAAACAGAAAGAAATTTAAAAA
AAAAA

FIGURE 125

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57689

<subunit 1 of 1, 406 aa, 1 stop

<MW: 46927, pI: 5.21, NX(S/T): 0

MHPAVFLSLPDLRCSLLLLVTWVFTPVTTEITSLATENIDEILNNADVALVNFYADWCRFSQ
MLHPIFEEASDVIKEEFPNENQVVFARVDCDQHSDIAQRYRISKYPTLKLFRNGMMMKREYR
GQRSVKALADYIRQQKSDPIQEIRDLAEITTLDRSKRNIIGYFEQXSDNYRVFERVANILH
DDCAFLSAFGDVSKPERYSGDNIIYKPPGHSAPDMVYLGAMTNFDVTYNWIQDKCVPLVREI
TFENGEELTEEGLPFLILFHKEDTESLEIFQNEVARQLISEKGTINFLHADCDKFRHPLLH
IQKTPADCFVIAIDSFRHMYVFGDFKDVLI PGKLKQFVFDLHSGKLHREFHHGPDPTDTAPG
EQAQDVASSPPESFQKLAPSEYRYTLLRDRDEL

Important features:

Signal peptide:

amino acids 1-29

Endoplasmic reticulum targeting sequence.

amino acids 403-406

Tyrosine kinase phosphorylation site.

amino acids 203-211

Thioredoxin family proteins

amino acids 50-66

FIGURE 126

ATTAAGGAAGAATTTCCAAATGAAAATCAAGTAGTNTTTGCCAGAGTNGATTGTGATCAGCA
CTCTGACATAGCCCAGAGATACAGGATAAGCAAATACCCAACCCCTCAAATTGTTTCGTAATG
GGATGATGATGAAGAGAGAATACAGGGGTCAGCGATCAGTGAAAGCATTTGGCAGATTA

10047194 102400
TGTGGT TGTGTT

FIGURE 127

AGAGGCCCTCTCTGGAAGTTGTCCCGGGTGTTCGCCGCNCGAGCCCGGGTCGAGAGGACNAGG
TGCCCGCTGCCTGGAGAATCCTCCGCTGCCGTCCGGCTCCCGGAGCCCAGCCCTTTCCTAACCC
AACCCAACTAGCCCNCTCCAGCCGCCAGCGCCTGTCCCTGTCTCNGGANCCAGCGTNACC
ATGCATCCTGCCGTCTTCTATCCTTACCCGACCTCAGATGCTCCCTTCTGCTCCTGGTAAC
TTGGGTTTTTACTCCTGTAACTGAAATAACNNGTCTTGATACNNAGAATATAGATGAAA
TTTTAAACNATGCTGATGTGGCTTTAGTCAATTTTATGCTGACTGGTGTCTTTTCAGTCAG
ATGTGGCATCCAATTTTGGAGANGCTTCCGATGTCATTAAGGAAGAATTTCCAAATGAAAA
TCAAGTAGTGTGTCAGAGTTGATTGTGATCAGCACTCTGACATAGCCCAGAGATACAGGA
TAAGCAAATACCCAACCTCAAATTGTTTCGTAATGGGATGATGATGAAGAGAGAATACAGG
GGTCAGCGATCAGTGAAAGCATTGGCAGATTACATCAGGC

1004764000

FIGURE 128

GCCCACGCGTCCGATGGCGTTTCACGTTTCGCGGCCCTTCTGCTACATGCTGGCGCTGCTGCTCA
CTGCCGCGCTCATCTTCTTCGCCATTTGGCACATTATAGCATTTGATGAGCTGAAGACTGAT
TACAAGAATCCTATAGACCAGTGAATACCCCTGAATCCCCCTTGTA

CTCCAGAGTACCTCAT
CCACGCTTTCCTCTGTGTCATGTTTCTTTGTGCAGCAGAGTGGCTTACACTGGGTCTCAATA
TGCCCCCTCTTGGCATATCATATTTGGAGGTATATGAGTAGACCAGTGATGAGTGGCCCAGGA
CTCTATGACCCACAAACCATCATGAATGCAGATATTCTAGCATATTGTCAGAAGGAAGGATG
GTGCAAATTAGCTTTTTATCTTCTAGCATTTTTTTACTACCTATATGGCATGATCTATGTTT
TGGTGAGCTCTTTAGAACAACACACAGAAGAAATTGGTCCAGTTAAGTGATGCAAAAAGCCAC
CAAATGAAGGGATTCTATCCAGCAAGATCCTGTCCAAGAGTAGCCTGTGGAATCTGATCAGT
TACTTTAAAAAATGACTCCTATTTTTTAAATGTTTCCACATTTTTGCTTGTGGAAGACTG
TTTTTCATATGTTATACTCAGATAAAGATTTTAAATGGTATTACGTATAAATTAATATAAAAT
GATTACCTCTGGTGTGACAGGTTTGAACCTGCACCTTCTTAAGGAACAGCCATAATCCTCTG
AATGATGCATTAATTACTGACTGTCTAGTACATTGGAAGCTTTTGTTTATAGGAACCTTGTA
GGGCTCAITTTGGTTTCATTGAAACAGTATCTAATTATAAATTAGCTGTAGATATCAGGTGC
TTCTGATGAAGTGAAAAATGTATATCTGACTAGTGGGAACTTCATGGGTTTCCTCATCTGTC
ATGTCGATGATTATATATGGATACATTACAAAAATAAAAAGCGGAAATTTTCCCTTCGCTT
GAATATTATCCCTGTATATTGCATGAATGAGAGATTTCCCATATTTCCATCAGAGTAATAAA
TATACTTGCTTTAATTCTTAAGCATAAGTAAACATGATATAAAAAATATATGCTGAATTACTT
GTGAAGAATGCATTTAAAGCTATTTTAAATGTGTTTTTATTTGTAAGACATTACTTATTAAG
AAATTGGTTATTATGCTTACTGTTCTAATCTGGTGGTAAAGGTATTCCTAAGAATTTCAGG
TACTACAGATTTTCAAACCTGAATGAGAGAAAAATTGTATAACCATCCTGCTGTTCCCTTAGT
GCAATACAATAAAACTCTGAAATTAAGACTC

3007491-16444

FIGURE 129

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA23330

<subunit 1 of 1, 144 aa, 1 stop

<MW: 16699, pI: 5.60, NX(S/T): 0

MAFTFAAFCYMLALLLTAALIFFAIWHIIAFDELKTDYKNPIDQCNTLNPLVLPEYLIHAFF
CVMFLCAAEWLTLGLNMPLLAYHIWRYMSRPVMSGPGLYDPTTIMNADILAYCQKEGWCKLA
FYLLAFFYYLYGMIYVLVSS

Important features:

Signal peptide:

amino acids 1-20

Type II transmembrane domain:

amino acids 11-31

Other transmembrane domain:

amino acids 57-77 and 123-143

FIGURE 130

ATTATAGCATTTGATGAGCTGAAGACTGATTACAAGATCCTATAGACCAGTGTAAATACCCCTG
AATCCCCCTGTACTCCCAGAGTACCTCATCCACGCTTTCTTCTGTGTCATGTTTCTTTGTGC
AGCAGAGTGGCTTACACTGGGTCTCAATATGCCCCCTCTTGGCATATCATATTTGGAGGTATA
TGAGTAGACCAGTGATGAGTGGCCCAGGACTCTATGACCCTACAACCATCATGAATGCAGAT
ATTCTAGCATATTGTCAGAAGGAAGGATGGTGCAAATTAGCTTTTATCTTCTAGCATTTTT
TTACTACCTATATGGCATGATCTATGTTTTGGTGAGCTCTTAGAACAACACACAGAAGAATT
GGTCCAGTTAAGTGCATGCAAAAAGCCACCAAATGAAGGGATTCTATCCAGCAAGATCCTGT
CCAAGAGTAGCCTGTGGAATCTGATCAGTTACTTTAAAAATG

100 200 300 400 500 600 700 800 900 1000

FIGURE 131

CGGACGCGTGGGGGAAACCTTCCGAGAAAAACAGCAACAAGCTGAGCTGCTGTGACAGAGGG
GAACAAGATGCGCGCGCCGAAGGGGAGCCTCTGGGTGAGGACCCAAGCTGGGGCTCCGCGCGC
TGCTGCTGCTGACCATGGCCTTGGCCGGAGGTTGGGGGACCGCTTCGGCTGAAGCATTTGAC
TCGGTCTTGGGTGATACGGCGTCTTGCCACCGGGCCTGTCAAGTTGACCTACCCCTTGACACAC
CTACCCCTAAGGAAGAGGAGTTGTACGCATGTCAGAGAGGTTGCAGGCTGTTTTCAATTTGTC
AGTTTGTGGATGATGGAATTGACTTAAATCGAACTAAATTGGAATGTGAATCTGCATGTACA
GAAGCATATCCCAATCTGATGAGCAATATGCTTGCCATCTTGGTTGCCAGAATCAGCTGCC
ATTCGCTGAAGTGAACAAGAACAACTTATGTCCTGATGCCAAAAATGCACCTACTCTTTC
CTCTAACTCTGGTGAGGTCATTCTGGAGTGACATGATGGAATCCGCACAGAGCTTCATAACC
TCTTCATGGAATTTTTATCTTCAAGCCGATGACGGAATAAGTTATATTCCAGTCTAAGCC
AGAAATCCAGTACGACCCACATTGGAGCAGGAGCCTACAAATTTGAGAGAATCATCTCTAA
GCAAAATGTCCTATCTGCAATGAGAAATCACAAAGCGCACAGGAATTTTCTGAAGATGGA
GAAAGTGATGGCTTTTTAAGATGCCTCTCTTAACTCTGGGTGGATTTTAACTACAACTCT
TGTCTCTCGGTGATGGTATTGCTTTGGATTGTTGTGCAACTGTTGCTACAGCTGTGGAGC
AGTATGTTCCCTCTGAGAAGCTGAGTATCTATGGTGACTTGGAGTTTATGAATGAACAAAAG
CTAAACAGATATCCAGCTTCTTCTCTTGTGGTTGTTAGATCTAAACTGAAGATCATGAAGA
AGCAGGGCCTCTACCTACAAAAGTGAATCTTGCTCATTCTGAAATTTAAGCATTCTTTCTTT
AAAAGACAAGTGAATAGACATCTAAATTCCTCCTCATAGAGCTTTTAAATGGTTTCA
TTGGATATAGGCCTTAAGAAATCACTATAAAATGCAATAAAGTTACTCAATCTGTG

FIGURE 132

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA26847

<subunit 1 of 1, 323 aa, 1 stop

<MW: 36223, pI: 5.06, NX(S/T): 1

MAAPKGS LWVRTQLGLPPLLLLT MALAGSGTASAEAFDSVLGDTASCHRA CQLTYPLHTYP
KEEELYACQRGCRLFSICQFVDDGIDLNR TKLECESACTEAYSQSDEQYACHLG CQNQLPFA
ELRQEQLM SLMPKMHLLFPLTLVRSFWS DMMDSAQSFITSSWTFY LQADDGKIVIFQSKPEI
QYAPHLEQEPTNLRESSLSKMSY LQMRNSQAHRNFLEDGESD GFLRCLSLNSGWILTTTLVL
SVMVLLWICCATVATAVEQYVPSEKLSIYGDLEFMNEQKLNRYPASSLVVVRSKTEDHEEAG
PLPTKVNLAHSEI

Important features:

Signal peptide:

amino acids 1-31

Transmembrane domain:

amino acids 241-260

N-glycosylation site.

amino acids 90-93

FIGURE 133

TTGGGTGATACGGCGTCTTGCCACCGGGCCTGTCAGTTGACCTACCCCTTGCACACCTACCC
TAAGGAAGAGGAGTTGTACGCATGTCAGAGAGGTTGCAGGCTGTTTCAATTGTCAGTTTG
TGGATGATGGAATTGACTTAAATCGAACTAAATTGGAATGTGAATCTGCATGTACAGAAGCA
TATTCCCAATCTGATGAGCAATATGCTTGCCATCTTGGTTGCCAGAATCAGCTGCCATTCCG
TGAAGTGAACAAGAACAACCTTATGTCCCTGATGCCAAAAATGCACCTACTCTTTCCTCTAA
CTCTGGTGAGGTCATTCTGGAGTGACATGATGGACTCCGC

100123456789101112131415161718192021222324252627282930313233343536373839404142434445464748495051525354555657585960616263646566676869707172737475767778798081828384858687888990919293949596979899100

Figure 1. The 12 test items of the T-LESQ.

CACACTGGCCGGGATCTTTTAGAGTCTTTTGACCTTGACCAAGGGTCNGGAAAAACAGCAACA
GCTGAGCTGCTGTGACAGAGGGAACAAGATGGCGGCGCCGAAGGAGCCTTTGGGTGAGGAC
CCAACCTGGGGCTCCCGCCGCTGCTGCTGTGACCATGGCCTTGGCCGGAGGTTGCGGGACCG
CTTCGGCTGAAGATTGACTCGGTCTTGGGTGATACGGCGTCTTGCCACCGGGCCTGTGAG
TTGACCTACCCCTTGACACCTACCTTAAGGAAGAGGAGTTGTACGCATGTCAGAGAGGTTG
CAGGCTGTTTTCAATTGTGACGTTTGTGGATGATGGAATTGACTTAAATCGAACTAAATTGG
AATGTGAATCTGCATGTACAGAAGCATATTTCCCAATCTGATGAGCAATATGCTTGCCATCTT
GGTTGCCAGAATCAGCTGCCATTGCTGAACTGAGACAAGAACAACCTTATGTCCCTGATGCC
AAAAATGCACCTACTCTTTCCTCTAACTCTGGTGAGGTCATTCTGGAGTGACATGATGGACT
CCGC

THE

Abstract

FIGURE 136

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53974

<subunit 1 of 1, 468 aa, 1 stop

<MW: 54393, pI: 5.63, NX(S/T): 2

MGRGWGFLFGLLGAVWLLSSGHGEEQPPETAAQRCFCQVSGYLDDCTCDVETIDRFNNYRLF
PRLQKLLSDYFRYYKVNLRPCFFWNDISQCGRRDCAVKPCQSDEVDPGIGKSASYKYSEEA
NNLIBECEQAERLGAVDESLSEETQKAVLQWTKHDDSSDNFCEADDIQSPEAEYVDLLNPE
RYTGYKGPDAWKIWNVIYEENCFKPQTIKRPLNPLASGQGTSEENTFYSWLEGLCVEKRAFY
RLISGLHASINVHLSARYLLQETWLEKKWGHNITEFQQRFDGILTEGEGPRRLKNLYFLYLI
ELRALSKVLPPFERPDFQLFTGNKIQDEENKMLLLEILHEIKSFPLHFDENSFFAGDKKEAH
KLKEDFRLHFRNISRIMDCVGCFCRLWGKLQTQGLGTALKILFSEKLIANMPESGSPSYEFH
LTRQEIVSLFNAFGRISTSVKELENFRNLLQNIH

Important features:

Signal peptide:

amino acids 1-23

N-glycosylation site.

amino acids 280-283 and 384-387

Amidation site.

amino acids 94-97

Glycosaminoglycan attachment site.

amino acids 20-23 and 223-226

Aminotransferases class-V pyridoxal-phosphate

amino acids 216-222

Interleukin-7 proteins

amino acids 338-343

GCTGGAAATATGGATGTCATCTACGAGAAACTGTTTTAAGCCACAGACAATTAAAAGACCTT
TAAATCCCTTTGGCTTCTGGTCAAGGGACAAGTGAAGAGNACCTTTTACAGTTGGCTAGAA
GGTCTCTGTGTAGAAAAAAGAGCATTCTACAGACTTATATCTGGCCTACATGCAAGCATTAA
TGTGCATTTGAGTGC AAGATATCTTTTACAAGAGACCTGGTTAGAAAAGAAATGGGGACACA
ACATTACAGAATTTNAACAGCGAATTTGATGGAATTTTGACTGAAGGAGAAGGTCCAAGAAGG
CTTAAGAACTTGATTTTTCTCTACTTAATAGAACTAAGGGCTTTATCCAAAGTGTTACCATT
CTTNGAGCGCCACAGATTTTCAACTNNTTACTGGAATAAAATTCAGGATGAGGNAACAAAA
TGTTACTTTTGGAAATACTTCATGAAATCAAGTCATTTTCCTTTGCATTTTGATGAGAAATCA
TTTTTTTGCTG

FIGURE 138

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGTGGGAGGGGGCAGGATGGGAGGGAA
AGTGAAGAAAAACAGAAAAGGAGAGGGACAGAGGCCAGAGGACTTCTCATACTGGACAGAAAC
CGATCAGGCATGGAACTCCCCTTCGTCACTCACCTGTTCTTGCCCTGGTGTTCCTGACAGG
TCTCTGCTCCCCCTTAACTTGGATGAACATCACCCACGCCTATTCCCAGGGCCACCAGAAAG
CTGAATTGGGATACAGTGTCTTACAACATGTTGGGGGTGGACAGCGATGGATGCTGGTGGGC
GCCCCCTGGGATGGGCCTTCAGGCGACCGGAGGGGGGACGTTTATCGCTGCCCTGTAGGGGG
GGCCCACAATGCCCATGTGCGCAAGGGCCACTTAGGTGACTACCAACTGGGAAATTCATCTC
ATCCTGCTGTGAATATGCACTGGGGATGTCTCTGTTAGAGACAGATGGTGATGGGGGAATTC
ATGGTGAGCTAAGGAGAGGGTGGTGGCAGTGTCTCTGAAGGTCCATAAAAGAAAAAGAGAA
GTGTGGTAAGGGAAAAATGGTCTGTGTGGAGGGGTCAAGGAGTAAAAAACCTAGAAAGCAAA
AGGTAGGTAATGTCAGGGAGTAGTCTTCATGCCCTCCTTCAACTGGGAGCATGTTCTGAGGGT
GCCCTCCCAAGCCTGGGAGTAACATATTCCCCCATCCCCAGGCCTGTGCCCTCTCTGGTCT
CGTGCTTGTGGCAGCTCTGTCTTCAGTTCCTGGGATATGTGCCGTGTGGATGCTTCATTCCA
GCCTCAGGGAAGCCTGGCACCCACTGCCCAACGTGAGCCAGAGGAAGGCTGAGTACTTGGTT
CCCAGAAGGAGATACCTGGGTGGGAAAAAGATGGGGCAAAGCGGTATGATGCCTGGCAAAGGG
CCTGCATGGCTATCCTCATTGCTACCTAATGTGCTTGCAAAGCTCCATGTTTCCTAACAGA
TTCAGACTCCTGGCCAGGTGTGGTGGCCACACCTGTAATTCTAGCACTTTGGGAGGCCAAG
GTGGGCAGATCACTTGAGGTGAGGAGTTCAAGACCAGCCTGGCCAACTGGTGAAACTCCAT
CTCTACTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA
ATCTACTCGGAGGCTAAGACAGGAGACTCTCACTTCAACCCAGGAGGTGGAGGTTGCGGTG
AGCCAAGATTGTGCCTCTGCACTCTAGCGTGGGTGACAGAGTAAGCGAGACTCCATCTCAAA
AATAATAATAATAATAATTAGACTCCTTATCAGGAGTCCATGATCTGGCCTGGCACAGTAA
CTCATGCCTGTAATCCCAACATTTTGGGAGGCCAACGAGGAGGATTGCTTGAGGTCTGGAG
GTTTGAGACCAGCCTGGGCAACATAGAAAGACCCCATCTCTAAATAAATGTTTTAAAAAT

FIGURE 139

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57039
><subunit 1 of 1, 124 aa, 1 stop
><MW: 13352, pI: 5.99, NX(S/T): 1
MELPFVTHLFLPLVFLTGLCSPFNLDEHHPRLFPGPPEAEFGYSVLQHVGGGQRWMLVGAPW
DGPSGDRRGDVYRCFVGGAHNAPCAKGHLGDYQLGNSSHPAVNMHLGMSLLETGDGGFMVS
```

Important features:

Signal peptide:

amino acids 1-22

Cell attachment sequence.

amino acids 70-73

N-glycosylation site.

amino acids 98-101

Integrins alpha chain proteins

amino acids 67-81

FIGURE 140

CACAGTTCCTCCACCATCACTCNTCCCATTCCTTCCAACCTTATTTTTAGCTTGCCATTGGGA
GGGGGCAGGATGGGAGGGGAAAGTGAAGAAAACAGAAAAGGAGAGGGACAGAGGCCAGAGGAC
TTCTCATACTGGACAGAAACCGATCAGGCATGGAACTCCCCCTTCGTCACTCACCTGTTCTTG
CCCCTGGTGTTCTTGACAGGTCTCTGCTCCCCCTTTAACCTGGATGAACATCACCCACGCCT
ATTCCAGGGCCACCAGAAGCTGAATTTGGATACAGTGTCTTACAACATGTTGGGGGTGGAC
AGCGATGGATGCTGGTGGGCGCCCCCTGGGATGGGCCTTCAGGCGACCGGAGGGGGACGTT
TATCGCTGCCCTGTAGGGGGGGCCCAATGCCCCATGTGCCAAGGGCCACTTAGGTGACTA
CCAACTGGGAAATTATCTCATCCTGCTGTGAATATGCACCTGGGGATGTCTCTGTTAGAGA
CAGATGGTGATGG

FIGURE 141

AAAGTTACATTTTCTCTGGAACCTCTCCTAGGCCACTCCCTGCTGATGCAACATCTGGGTTTG
GGCAGAAAGGAGGGTGCTTCGGAGCCCGCCCTTTCTGAGCTTCCTGGGCCGGCTCTAGAACA
ATTGAGGCTTCGCTCGGACTCAGACCTCAGCTCCAACATATGCATTCTGAAGAAAGATGGCT
GAGATGGACAGAATGCTTTATTTTGGAAAGAAACAATGTTCTAGGTCAAACCTGAGTCTACCA
AATGCGAGACTTTACAAATGGTTCTAGAAGAAATCTGGACAAGTCTTTTCATGTGGTTTTCT
ACGGCATTGATTCCATGTTTTGCTCACAGATGAAGTGGCCATTCTGCCTGCCCTCAGAACCTC
TCTGTACTCTCAACCAACATGAAGCATCTCTTGATGTGGAGCCCACTGATCGCGCTGGAGA
AACAGTGTACTATTCTGTCGAATACCAGGGGGAGTACGAGAGCCTGTACACGAGCCACATCT
GGATCCCCAGCAGCTGGTGCTCACTCACTGAAGTCTTGAGTGTGATGTCACTGATGACATC
ACGGCCACTGTGCCATACAACCTTCGTGTGTCAGGGCCACATTGGGCTCACAGACCTCAGCCTG
GAGCATCCTGAAGCATCCCTTTAATAGAAACTCAACCATCCTTACCCGACCTGGGATGGAGA
TCACCAAAGATGGCTTCCACCTGGTTATTGAGCTGGAGGACCTGGGGCCCCAGTTTGAGTTC
CTTGTCGCTACTGAGGAGGGGAGCCTGGTGCCGAGGAACATGTCAAATGGTGAGGAGTGG
GGGTATTCAGTGCACCTAGAAACCATGGAGCCAGGGGCTGCATACTGTGTGAAGGCCAGA
CATTGTCGAAGGCCATTGGGAGGTACAGCGCCTTCAGCCAGACAGAATGTGTGGAGTGCAA
GGAGAGGCCATTCCCTGGTACTGGCCCTGTTTGCCCTTGTGTGGCTTCATGCTGATCCTTGT
GGTCGTGCCACTGTTTCGTCTGGAATAATGGGCCGGCTGCTCCAGTACTCCTGTTGCCCGTGG
TGTTCCTCCAGACACCTTGAAAATAACCAATTACCCACAGAAGTTAATCAGCTGCAGAAGG
GAGGAGGTGGATGCCTGTGCCACGGCTGTGATGTCTCCTGAGGAACTCCTCAGGGCCTGGAT
CTC**ATAG**GTTTGCGGAAGGGCCAGGTGAAGCCGAGAACCTGGTCTGCATGACATGGAAGCC
ATGAGGGGACAAGTTGTGTTTTCTGTTTTCCGCCACGGACAAGGGATGAGAGAAGTAGGAAGA
GCCTGTTGTCTACAAGTCTAGAAGCAACCATCAGAGGCAGGGTGGTTTTGTCTAACGAAACAC
TGACTGAGGCTTAGGGGATGTGACCTCTAGACTGGGGGCTGCCACTTGCTGGCTGAGCAACC
CTGGGAAAAGTGACTTCATCCCTTCGGTCCTAAGTTTTCTCATCTGTAATGGGGGAATTACC
TACACACCTGTAAACACACACACAGAGTCTCTCTATATATACACACGTACACATAAA
TACCCAGCCTTGCAAGGCTAGAGGGAACTGGTGACACTCTACAGTCTGACTGATTTCAG
TGTTTCTGGAGAGCAGGACATAAATGTATGATGAGAATGATCAAGGACTCTACACACTGGGT
GGCTTGAGAGGCCCACTTTCCAGAATAATCCTTGAGAGAAAAGGAATCATGGGAGCAATGG
TGTTGAGTTCACCTCAAGCCCAATGCCGGTGTCAGAGGGGAATGGCTTAGCGAGCTCTACAGT
AGGTGACCTGGAGGAAGGTACAGCCCACTGAAAATGGGATGTGCATGAACACGGAGGATC
CATGAACACTACTGTAAAGTGTGTGACAGTGTGTGCACTGCAGACAGCAGGTGAAATGTATGT
GTGCAATGCGACGAGAATGCAGAAGTCAGTAAACATGTGCATGTTTGTGTGCTCCTTTTTTC
TGTTGGTAAAGTACAGAATTACAGCAAATAAAAAGGGCCACCTGGCCAAAAGCGGTAAAAAA
AAAAAAAAAA

1017291.102491

FIGURE 142

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57033

<subunit 1 of 1, 311 aa, 1 stop

<MW: 35076, pI: 5.04, NX(S/T): 2

MQTFTMVLEEIWTSFLFMWFFYALIPCLLTDEVAILPAPQNLSVLSTNMKHLMLMWSPVIAPGE
TVYYSVEYQGEYESLYTSHIWIPSSWCSLTEGPECVTDITATVPYNLVRATLGSQTSW
SILKHPFNRRNSTILTRPGMEITKDGPHLVIELEDLGPQFEFLVAYWRRREPGAEEHVKMVRSG
GIPVHLETMEPGAAYCVKAQTFVKAIGRYSAFSQTECVEVQGEAIPVLALFAFVGFMILILV
VVPLFVWKMGRLQLQYSCCPVVVLPDTLKITNSPQKLISCRREEVDACATAVMSPEELLRAWIS

Important features:

Signal peptide:

amino acids 1-29

Transmembrane domain:

amino acids 230-255

N-glycosylation site.

amino acids 40-43 and 134-137

Tissue factor proteins.

amino acids 92-119

Integrins alpha chain proteins

amino acids 232-262

FIGURE 143

TCCTGCTGATGCACATCTGGGTTTGGCAAAAGGAGGTTGCTTCGAGCCGCCCTTTCTAGCTT
CCTGGCCGGCTCTAGAACAATTCAAGGCTTCGCTGCGACTAGACCTCAGCTCCAACATATGCA
TTCTGAAGAAAAGATGGCTGAGATGACAGAATGCTTTATTTTGGAAAGAAACAATGTTCTAGG
TCAAACCTGAGTCTACCAAATGCAGACTTTCACAATGGTTCGAGAAGAAATCTGGACAAGTCT
TTTCATGTGGTTTTTCTACGCATTGATTCCATGTTTGCTCACAGATGAAGTGCCATTCTGC
CTGCCCCCTCAGAACCTCTCTGTACTCTCAACCAACATGAAGCATCTCTTGATGTGGAGCCCA
GTGATCGCGCCTGGAGAAAACAGTGTACTATCTGTTCGAATCCAGGGGGAGTACGAGAGCCT
GTACACGAGCCACATCTGGATCCCCAGCAGCTGGTGCTCACTCACTGAAGGTCCTGAGTGTG
ATGTCACTGATGACATCACGGCCACTGTGCCATACAACCTTTGTGTGAGGGCCACATTGGGC
TCACAGACCTCAGCCTGGAGCATCCTGAAGCATCCCTTTAATAGAACTCAACCATCCTTAC
CCGACCTGGGATGGAGATCACCAAAGATGGCTTNCACCTGGTTATTGAGCTGGAGGACCTGG
GGCCCCAGTTTGAGTTCCTTGTGGCCTANTGGAGGAGGGGCGAACCCCTTGCGGCGCAAGGG
GTTNGCGAACCCCTTGCGGCCGCTGGGGTATCTCTCGAGAAAAGAGAGGCCCAATATGACCC
ACATACTCAATATGGACGAANTGCTATTGTCCACCTGTTTGAGTGGCGCTGGGTGAT

FIGURE 144

CCCACGCGTCCGCCACGCGTCCGAGGGACAAGAGAGAAGAGAGACTGAAACAGGGAGAAGA
GGCAGGAGAGGAGGAGGTGGGGAGAGCACGAAGCTGGAGGCCGACACTGAGGGAGGGCGGGA
GGAGGTGAAGAAGGAGAGAGGGGAGAAGAGGCAGGAGCTGGAAGGAGAGAGGGAGGAGGAG
GAGGAGATGCGGGATGGAGACCTGGAGTTAGGTGGCTTGGGAGAGCTTAATGAAAAGAGAAC
GGAGAGGAGGTGTGGGTTAGGAACCAAGAGGTAGCCCTGTGGGCAGCAGAAGGCTGAGAGGA
GTAGGAAGATCAGGAGCTAGAGGGAGACTGGAGGGTTCCGGGAAAAGAGCAGAGGAAAGAGG
AAAGACACAGAGAGACGGGAGAGAGAAGAAGAGTGGGTTTGAAGGGCGGATCTCAGTCCCTG
GCTGCTTTGGCATTGGGGAAGTGGGACTCCCTGTGGGGAGGAGGAAAGCTGGAAGTCTT
GGAGGGACAGGGTCCCAGAAGGAGGGGACAGAGGAGCTGAGAGAGGGGGCAGGGCGTTGGG
CAGGGGTCCCTCGGAGGCCTCCTGGGGATGGGGGCTGCAGCTCGTCTGAGCGCCCCGAGC
GCTGGTACTCTGGGCTGCACTGGGGGAGCAGCTCACATCGGACCAGCACCTGACCCCGAGG
ACTGGTGGAGCTACAAGGATAATCTCCAGGGAAGTTCGTGCCAGGGCCTCCTTTCTGGGGC
CTGGTGAATGCAGCGTGGAGTCTGTGTGCTGTGGGAAGCGGCAGAGCCCCGTGGATGTGGA
GCTGAAGAGGGTTCTTTATGACCCCTTTCTGCCCCATTAAAGGCTCAGCACTGGAGGAGAGA
AGCTCCGGGGAACCTTGTACAACACCGGCCGACATGTCTCCTTCTGCTGCACCCCGACCT
GTGGTCAATGTGTCTGGAGGTCCCTCCTTTACAGCCACCGACTCAGTGAAGTGCGGCTGCT
GTTTGGAGCTCGCGACGGAGCCGGCTCGGAACATCAGATCAACCACAGGGCTTCTCTGCTG
AGGTGCAGCTCATTCACTTCAACCAGGAAGTCTACGGGAATTTAGCGCTGCCTCCCGCGGC
CCCAATGGCCTGGCCATTCTCAGCCTCTTTGTCAACGTTGCCAGTACCTCTAACCATTCTCT
CAGTCGCCTCCTTAACCGGACACCATCACTCGCATCTCCTCAAGAATGATGCCCTACTTTC
TTCAAGACCTGAGCCTGGAGCTCCTGTTCCCTGAATCCTTCGGCTTCATCACCTATCAGGGC
TCTCTCAGCACCCCGCCCTGCTCCGAGACTGTACCTGGATCCTCATTGACCGGGCCCTCAA
TATCACCTCCCTTACATGCACTCCCTGAGACTCCTGAGCCAGAATCCTCCATCTCAGATCT
TCCAGAGCCTCAGCGGTAAACAGCCGGCCCTGCAGCCCTTGGCCACAGGGCACTGAGGGGC
AACAGGGACCCCCGACCCCGAGAGCGCTGCCGAGGCCCAACTACCGCTGCATGTGGA
TGGTGTCCCCATGGTCTGCTGAGACTCCCTTCGAGGATTGCACCCGCCCGTCTTAAGCCTC
CCCAAGGCGAGGGGAGTTACCCCTAAAACAAAGCTATTAAAGGGACAGAATACTTA

FIGURE 145

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA34353
<subunit 1 of 1, 328 aa, 1 stop
<MW: 36238, pI: 9.90, NX(S/T): 3
MGAAARLSAPRALVLWAAALGAAAHIGPAPDPEDWWSYKDNLQGNFVPGPPFWGLVNAAWSLC
AVGKRQSPVDVELKRVLYDPFLPPLRLSTGGEKLRGTLTYNTGRHVSFLLPAPRPVVNVSGGPL
LYSHRLESELRLFLGARDGAGSEHQINHQGFSAEVQLIHFNQELYGNFSAASRGPNGLAILSL
FVNVASTSNPFLSRLLNRDTITRISYKNDAYFLQDLSELELFPESFGFITYQGSLSLTPPCSE
TVTWILIDRALNITSLQMHSRLRLSQNPFSQIFQSLSGNSRPLQPLAHRALRGNRDPRHPER
RCRGPNYRLHVDGVPHGR
```

Important features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 177-199

N-glycosylation site.

amino acids 118-121, 170-173 and 260-263

Eukaryotic-type carbonic anhydrases proteins

amino acids 222-270, 128-164 and 45-92

FIGURE 147

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45417

<subunit 1 of 1, 500 aa, 1 stop

<MW: 56888, pI: 8.53, NX(S/T): 2

MKCTAREWLRVTTVLFMARAI PAMVVPNATLLEKLLEKYMDEEDGEWWIAKQGRKRAITDNDM
QSI L D L H N K L R S Q V Y P T A S N M E Y M T W D V E L E R S A E S W A E S C L W E H G P A S L L P S I G Q N L G A H W
G R Y R P P T F H V Q S W Y D E V K D F S Y P Y E H E C N P Y C P F R C S G P V C T H Y T Q V V W A T S N R I G C A I N L C
H N M N I W G Q I W F K A V Y L V C N Y S P K G N W W G H A P Y K H G R P C S A C P P S F G G C R E N L C Y K E G S D R Y
Y P F R E E T N E I E R Q Q S Q V H D T H V R T R S D D S S R N E V I S A Q Q M S Q I V S C E V R L R D Q C K G T T C N R
Y E C P A G C L D S K A K V I G S V H Y E M Q S S I C R A A I H Y G I I D N D G G W V D I T R Q G R K H Y F I K S N R N G I
Q T I G K Y Q S A N S F T V S K V T Q A V T C E T T V E Q L C P F H K P A S H C P R V Y C P R N C M Q A N P H Y A R V I G
T R V Y S D L S S I C R A A V H A G V V R N H G G Y V D V M P V D K R K T Y I A S F Q N G I F S E S L Q N P P G G K A F R V
FAVV

Important features:

Signal peptide:

amino acids 1-20

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 protein

amino acids 165-186, 196-218, 134-146, 96-108 and 58-77

N-glycosylation site

amino acids 28-31

FIGURE 148

GCGGAGACAAGCGCAGAGCGCAGCGCACGGCCACAGACAGCCCTGGGCATCCACCGACGGCG
CAGCCGGAGCCAGCAGAGCCGGAAGCGCGCCCCGGGCAGAGAAAGCCGAGCAGAGCTGGGT
GGCGTCTCCGGGCCCGGCTCCGACGGGCCAGCGCCCTCCCCATGTCCCTGCTCCACGCGCG
CGCCCCCTCCGGTCAGCATGAGGCTCCTGGCGGCCGCGCTGCTCCTGCTGCTGCTGGCGCTGT
ACACCGCGCGTGTGGACGGGTCCAAATGCAAAGTGTCTCCCGGAAGGGACCCAAGATCCGCTAC
AGCGAGTGAAGAAGCTGGAATGAAGCCAAAGTACCGGCACTGCGAGGAGAAGATGGTTAT
CATCACCAACAGAGCGTGTCCAGGTACCGAGGTCAGGAGCACTGCCTGCAACCCCAAGCTGC
AGAGCACCAAGCGCTTCATCAAGTGGTACAACGCCTGGAACGAGAAGCGAGGGTCTACGAA
GAATAGGGTGAAAAAACCTCAGAAGGGAAAACTCCAAACCAAGTTGGGAGACTTGTGCAAGGA
CTTTGCAGATTAAAAAAGGCTTTTGTGCAAGGCTTTTGTGCAAGGCTTTTGTGCAAGGCTTTT
TTTCTCAGGCATAAGACACAAATTATATATTGTTATGAAGCACTTTTTTACCAACGGTCAG
TTTTTACATTTTATAGCTGCGTGCGAAAGGCTTCCAGATGGGAGACCCATCTCTCTTGTGCT
CCAGACTTCATCAGAGGTGCTTTTTATCAAAAAGGGGAAAACTCATGCTTTTCTTTTAA
AAAATGCTTTTTTGTATTGTCCATACGTCACTATACATCTGAGCTTTATAAGCGCCCGGA
GGAACAATGAGCTTGGTGGACACATTCATTGCAGTGTGTGCTCCATTCTAGCTTGGGAAGC
TTCCGCTTAGAGGTCTCGCGCCTCGGCACAGCTGCCACGGGCTCTCCTGGGCTTATGGCCG
GTCACAGCCTCAGTGTGACTCCACAGTGGCCCCGTGAGCCGGCAAGCAGGAGCAGGTCTCT
CTGCATCTGTCTCTGAGGAACCTCAAGTTGGTTGCCAGAAAAATGTGCTTCATTCCCCCT
GGTTAATTTTACACACCCTAGGAAACATTTCCAAGATCCTGTGATGGCGAGACAAATGATC
CTTAAAGAAGGTGTGGGGTCTTTCCCAACCTGAGGAITTTCTGAAAGGTTACAGGTTCAATA
TTTAATGCTTCAGAAGCATGTGAGGTTCCCAACACTGTGAGCAAAAACCTTAGGAGAAAACT
TAAAAATATATGAATACATGCGCAATACACAGCTACAGACACACATTCTGTTGACAAGGGAA
AACCTTCAAAGCATGTTTCTTTCCCTCACCACAACAGAACATGCAGTACTAAGCAATATAT
TTGTGATTTCCCATGTAATTCTTCAATGTTAAACAGTGCAGTCTCTTTTGGAAAGCTAAGAT
GACCATGCGCCCTTTCCTCTGTACATATACCCTTAAGAACGCCCCCTCCACACACTGCCCCC
CAGTATATGCCGATTTGACTGCTGTGTATATGCTATGTACATGTGAGAAACCATTAGCAT
TGCATGCAGGTTTCATATTCTTTCTAAGATGGAAAGTAATAAAATATATTTGAAATGTAAAA
AAAAAAAAAA

FIGURE 149

MSLLPRRAPPVSMRLAAALLLLLLLALYTARVDGSKCKCSRKGPKIRYSDVKKLEMKPKYPH
CEEKMVIITTKSVSRYRGQEHCLHPKLQSTKRFIKWYNWNEKRRVYEE

Signal sequence:

amino acids 1-34

1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051 1052 1053 1054 1055 1056 1057 1058 1059 1060 1061 1062 1063 1064 1065 1066 1067 1068 1069 1070 1071 1072 1073 1074 1075 1076 1077 1078 1079 1080 1081 1082 1083 1084 1085 1086 1087 1088 1089 1090 1091 1092 1093 1094 1095 1096 1097 1098 1099 1100 1101 1102 1103 1104 1105 1106 1107 1108 1109 1110 1111 1112 1113 1114 1115 1116 1117 1118 1119 1120 1121 1122 1123 1124 1125 1126 1127 1128 1129 1130 1131 1132 1133 1134 1135 1136 1137 1138 1139 1140 1141 1142 1143 1144 1145 1146 1147 1148 1149 1150 1151 1152 1153 1154 1155 1156 1157 1158 1159 1160 1161 1162 1163 1164 1165 1166 1167 1168 1169 1170 1171 1172 1173 1174 1175 1176 1177 1178 1179 1180 1181 1182 1183 1184 1185 1186 1187 1188 1189 1190 1191 1192 1193 1194 1195 1196 1197 1198 1199 1200 1201 1202 1203 1204 1205 1206 1207 1208 1209 1210 1211 1212 1213 1214 1215 1216 1217 1218 1219 1220 1221 1222 1223 1224 1225 1226 1227 1228 1229 1230 1231 1232 1233 1234 1235 1236 1237 1238 1239 1240 1241 1242 1243 1244 1245 1246 1247 1248 1249 1250 1251 1252 1253 1254 1255 1256 1257 1258 1259 1260 1261 1262 1263 1264 1265 1266 1267 1268 1269 1270 1271 1272 1273 1274 1275 1276 1277 1278 1279 1280 1281 1282 1283 1284 1285 1286 1287 1288 1289 1290 1291 1292 1293 1294 1295 1296 1297 1298 1299 1300 1301 1302 1303 1304 1305 1306 1307 1308 1309 1310 1311 1312 1313 1314 1315 1316 1317 1318 1319 1320 1321 1322 1323 1324 1325 1326 1327 1328 1329 1330 1331 1332 1333 1334 1335 1336 1337 1338 1339 1340 1341 1342 1343 1344 1345 1346 1347 1348 1349 1350 1351 1352 1353 1354 1355 1356 1357 1358 1359 1360 1361 1362 1363 1364 1365 1366 1367 1368 1369 1370 1371 1372 1373 1374 1375 1376 1377 1378 1379 1380 1381 1382 1383 1384 1385 1386 1387 1388 1389 1390 1391 1392 1393 1394 1395 1396 1397 1398 1399 1400 1401 1402 1403 1404 1405 1406 1407 1408 1409 1410 1411 1412 1413 1414 1415 1416 1417 1418 1419 1420 1421 1422 1423 1424 1425 1426 1427 1428 1429 1430 1431 1432 1433 1434 1435 1436 1437 1438 1439 1440 1441 1442 1443 1444 1445 1446 1447 1448 1449 1450 1451 1452 1453 1454 1455 1456 1457 1458 1459 1460 1461 1462 1463 1464 1465 1466 1467 1468 1469 1470 1471 1472 1473 1474 1475 1476 1477 1478 1479 1480 1481 1482 1483 1484 1485 1486 1487 1488 1489 1490 1491 1492 1493 1494 1495 1496 1497 1498 1499 1500 1501 1502 1503 1504 1505 1506 1507 1508 1509 1510 1511 1512 1513 1514 1515 1516 1517 1518 1519 1520 1521 1522 1523 1524 1525 1526 1527 1528 1529 1530 1531 1532 1533 1534 1535 1536 1537 1538 1539 1540 1541 1542 1543 1544 1545 1546 1547 1548 1549 1550 1551 1552 1553 1554 1555 1556 1557 1558 1559 1560 1561 1562 1563 1564 1565 1566 1567 1568 1569 1570 1571 1572 1573 1574 1575 1576 1577 1578 1579 1580 1581 1582 1583 1584 1585 1586 1587 1588 1589 1590 1591 1592 1593 1594 1595 1596 1597 1598 1599 1600 1601 1602 1603 1604 1605 1606 1607 1608 1609 1610 1611 1612 1613 1614 1615 1616 1617 1618 1619 1620 1621 1622 1623 1624 1625 1626 1627 1628 1629 1630 1631 1632 1633 1634 1635 1636 1637 1638 1639 1640 1641 1642 1643 1644 1645 1646 1647 1648 1649 1650 1651 1652 1653 1654 1655 1656 1657 1658 1659 1660 1661 1662 1663 1664 1665 1666 1667 1668 1669 1670 1671 1672 1673 1674 1675 1676 1677 1678 1679 1680 1681 1682 1683 1684 1685 1686 1687 1688 1689 1690 1691 1692 1693 1694 1695 1696 1697 1698 1699 1700 1701 1702 1703 1704 1705 1706 1707 1708 1709 1710 1711 1712 1713 1714 1715 1716 1717 1718 1719 1720 1721 1722 1723 1724 1725 1726 1727 1728 1729 1730 1731 1732 1733 1734 1735 1736 1737 1738 1739 1740 1741 1742 1743 1744 1745 1746 1747 1748 1749 1750 1751 1752 1753 1754 1755 1756 1757 1758 1759 1760 1761 1762 1763 1764 1765 1766 1767 1768 1769 1770 1771 1772 1773 1774 1775 1776 1777 1778 1779 1780 1781 1782 1783 1784 1785 1786 1787 1788 1789 1790 1791 1792 1793 1794 1795 1796 1797 1798 1799 1800 1801 1802 1803 1804 1805 1806 1807 1808 1809 1810 1811 1812 1813 1814 1815 1816 1817 1818 1819 1820 1821 1822 1823 1824 1825 1826 1827 1828 1829 1830 1831 1832 1833 1834 1835 1836 1837 1838 1839 1840 1841 1842 1843 1844 1845 1846 1847 1848 1849 1850 1851 1852 1853 1854 1855 1856 1857 1858 1859 1860 1861 1862 1863 1864 1865 1866 1867 1868 1869 1870 1871 1872 1873 1874 1875 1876 1877 1878 1879 1880 1881 1882 1883 1884 1885 1886 1887 1888 1889 1890 1891 1892 1893 1894 1895 1896 1897 1898 1899 1900 1901 1902 1903 1904 1905 1906 1907 1908 1909 1910 1911 1912 1913 1914 1915 1916 1917 1918 1919 1920 1921 1922 1923 1924 1925 1926 1927 1928 1929 1930 1931 1932 1933 1934 1935 1936 1937 1938 1939 1940 1941 1942 1943 1944 1945 1946 1947 1948 1949 1950 1951 1952 1953 1954 1955 1956 1957 1958 1959 1960 1961 1962 1963 1964 1965 1966 1967 1968 1969 1970 1971 1972 1973 1974 1975 1976 1977 1978 1979 1980 1981 1982 1983 1984 1985 1986 1987 1988 1989 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020 2021 2022 2023 2024 2025 2026 2027 2028 2029 2030 2031 2032 2033 2034 2035 2036 2037 2038 2039 2040 2041 2042 2043 2044 2045 2046 2047 2048 2049 2050 2051 2052 2053 2054 2055 2056 2057 2058 2059 2060 2061 2062 2063 2064 2065 2066 2067 2068 2069 2070 2071 2072 2073 2074 2075 2076 2077 2078 2079 2080 2081 2082 2083 2084 2085 2086 2087 2088 2089 2090 2091 2092 2093 2094 2095 2096 2097 2098 2099 2100 2101 2102 2103 2104 2105 2106 2107 2108 2109 2110 2111 2112 2113 2114 2115 2116 2117 2118 2119 2120 2121 2122 2123 2124 2125 2126 2127 2128 2129 2130 2131 2132 2133 2134 2135 2136 2137 2138 2139 2140 2141 2142 2143 2144 2145 2146 2147 2148 2149 2150 2151 2152 2153 2154 2155 2156 2157 2158 2159 2160 2161 2162 2163 2164 2165 2166 2167 2168 2169 2170 2171 2172 2173 2174 2175 2176 2177 2178 2179 2180 2181 2182 2183 2184 2185 2186 2187 2188 2189 2190 2191 2192 2193 2194 2195 2196 2197 2198 2199 2200 2201 2202 2203 2204 2205 2206 2207 2208 2209 2210 2211 2212 2213 2214 2215 2216 2217 2218 2219 2220 2221 2222 2223 2224 2225 2226 2227 2228 2229 2230 2231 2232 2233 2234 2235 2236 2237 2238 2239 2240 2241 2242 2243 2244 2245 2246 2247 2248 2249 2250 2251 2252 2253 2254 2255 2256 2257 2258 2259 2260 2261 2262 2263 2264 2265 2266 2267 2268 2269 2270 2271 2272 2273 2274 2275 2276 2277 2278 2279 2280 2281 2282 2283 2284 2285 2286 2287 2288 2289 2290 2291 2292 2293 2294 2295 2296 2297 2298 2299 2300 2301 2302 2303 2304 2305 2306 2307 2308 2309 2310 2311 2312 2313 2314 2315 2316 2317 2318 2319 2320 2321 2322 2323 2324 2325 2326 2327 2328 2329 2330 2331 2332 2333 2334 2335 2336 2337 2338 2339 2340 2341 2342 2343 2344 2345 2346 2347 2348 2349 2350 2351 2352 2353 2354 2355 2356 2357 2358 2359 2360 2361 2362 2363 2364 2365 2366 2367 2368 2369 2370 2371 2372 2373 2374 2375 2376 2377 2378 2379 2380 2381 2382 2383 2384 2385 2386 2387 2388 2389 2390 2391 2392 2393 2394 2395 2396 2397 2398 2399 2400 2401 2402 2403 2404 2405 2406 2407 2408 2409 2410 2411 2412 2413 2414 2415 2416 2417 2418 2419 2420 2421 2422 2423 2424 2425 2426 2427 2428 2429 2430 2431 2432 2433 2434 2435 2436 2437 2438 2439 2440 2441 2442 2443 2444 2445 2446 2447 2448 2449 2450 2451 2452 2453 2454 2455 2456 2457 2458 2459 2460 2461 2462 2463 2464 2465 2466 2467 2468 2469 2470 2471 2472 2473 2474 2475 2476 2477 2478 2479 2480 2481 2482 2483 2484 2485 2486 2487 2488 2489 2490 2491 2492 2493 2494 2495 2496 2497 2498 2499 2500 2501 2502 2503 2504 2505 2506 2507 2508 2509 2510 2511 2512 2513 2514 2515 2516 2517 2518 2519 2520 2521 2522 2523 2524 2525 2526 2527 2528 2529 2530 2531 2532 2533 2534 2535 2536 2537 2538 2539 2540 2541 2542 2543 2544 2545 2546 2547 2548 2549 2550 2551 2552 2553 2554 2555 2556 2557 2558 2559 2560 2561 2562 2563 2564 2565 2566 2567 2568 2569 2570 2571 2572 2573 2574 2575 2576 2577 2578 2579 2580 2581 2582 2583 2584 2585 2586 2587 2588 2589 2590 2591 2592 2593 2594 2595 2596 2597 2598 2599 2600 2601 2602 2603 2604 2605 2606 2607 2608 2609 2610 2611 2612 2613 2614 2615 2616 2617 2618 2619 2620 2621 2622 2623 2624 2625 2626 2627 2628 2629 2630 2631 2632 2633 2634 2635 2636 2637 2638 2639 2640 2641 2642 2643 2644 2645 2646 2647 2648 2649 2650 2651 2652 2653 2654 2655 2656 2657 2658 2659 2660 2661 2662 2663 2664 2665 2666 2667 2668 2669 2670 2671 2672 2673 2674 2675 2676 2677 2678 2679 2680 2681 2682 2683 2684 2685 2686 2687 2688 2689 2690 2691 2692 2693 2694 2695 2696 2697 2698 2699 2700 2701 2702 2703 2704 2705 2706 2707 2708 2709 2710 2711 2712 2713 2714 2715 2716 2717 2718 2719 2720 2721 2722 2723 2724 2725 2726 2727 2728 2729 2730 2731 2732 2733 2734 2735 2736 2737 2738 2739 2740 2741 2742 2743 2744 2745 2746 2747 2748 2749 2750 2751 2752 2753 2754 2755 2756 2757 2758 2759 2760 2761 2762 2763 2764 2765 2766 2767 2768 2769 2770 2771 2772 2773 2774 2775 2776 2777 2778 2779 2780 2781 2782 2783 2784 2785 2786 2787 2788 2789 2790 2791 2792 2793 2794 2795 2796 2797 2798 2799 2800 2801 2802 2803 2804 2805 2806 2807 2808 2809 2810 2811 2812 2813 2814 2815 2816 2817 2818 2819 2820 2821 2822 2823 2824 2825 2826 2827 2828 2829 2830 2831 2832 2833 2834 2835 2836 2837 2838 2839 2840 2841 2842 2843 2844 2845 2846 2847 2848 2849 2850 2851 2852 2853 2854 2855 2856 2857 2858 2859 2860 2861 2862 2863 2864 2865 2866 2867 2868 2869 2870 2871 2872 2873 2874 2875 2876 2877 2878 2879 2880 2881 2882 2883 2884 2885 2886 2887 2888 2889 2890 2891 2892 2893 2894 2895 2896 2897 2898 2899 2900 2901 2902 2903 2904 2905 2906 2907 2908 2909 2910 2911 2912 2913 2914 2915 2916 2917 2918 2919 2920 2921 2922 2923 2924 2925 2926 2927 2928 2929 2930 2931 2932 2933 2934 2935 2936 2937 2938 2939 2940 2941 2942 2943 2944 2945 2946 2947 2948 2949 2950 2951 2952 2953 2954 2955 2956 2957 2958 2959 2960 2961 2962 2963 2964 2965 2966 2967 2968 2969 2970 2971 2972 2973 2974 2975 2976 2977 2978 2979 2980 2981 2982 2983 2984 2985 2986 2987 2988 2989 2990 2991 2992 2993 2994 2995 2996 2997 2998 2999 3000 3001 3002 3003 3004 3005 3006 3007 3008 3009 3010 3011 3012 3013 3014 3015 3016 3017 3018 3019 3020 3021 3022 3023 3024 3025 3026 3027 3028 3029 3030 3031 3032 3033 3034 3035 3036 3037 3038 3039 3040 3041 3042 3043 3044 3045 3046 3047 3048 3049 3050 3051 3052 3053 3054 3055 3056 3057 3058 3059 3060 3061 3062 3063 3064 3065 3066 3067 3068 3069 3070 3071 3072 3073 3074 3075 3076 3077 3078 3079 3080 3081 3082 3083 3084 3085 3086 3087 3088 3089 3090 3091 3092 3093 3094 3095 3096 3097 3098 3099 3100 3101 3102 3103 3104 3105 3106 3107 3108 3109 3110 3111 3112 3113 3114 3115 3116 3117 3118 3119 3120 3121 3122 3123 3124 3125 3126 3127 3128 3129 3130 3131 3132 3133 3134 3135 3136 3137 3138 3139 3140 3141 3142 3143 3144 3145 3146 3147 3148 3149 3150 3151 3152 3153 3154 3155 3156 3157 3158 3159 3160 3161 3162 3163 3164 3165 3166 3167 3168 3169 3170 3171 3172 3173 3174 3175 3176 3177 3178 3179 3180 3181 3182 3183 3184 3185 3186 3187 3188 3189 3190 3191 3192 3193 3194 3195 3196 3197 3198 3199 3200 3201 3202 3203 3204 3205 3206 3207 3208 3209 3210 3211 3212 3213 3214 3215 3216 3217 3218 3219 3220 3221 3222 3223 3224 3225 3226 3227 3228 3229 3230 3231 3232 3233 3234 3235 3236 3237 3238 3239 3240 3241 3242 3243 3244 3245 3246 3247 3248 3249 3250 3251 3252 3253 3254 3255 3256 3257 3258 3259 3260 3261 3262 3263 3264 3265 3266 3267 3268 3269 3270 3271 3272 3273 3274 3275 3276 3277 3278 3279 3280 3281 3282 3283 3284 3285 3286 3287 3288 3289 3290 3291 3292 3293 3294 3295 3296 3297 3298 3299 3300 3301 3302 3303 3304 3305 3306 3307 3308 3309 3310 3311 3312 3313 3314 3315 3316 3317 3318 3319 3320 3321 3322 3323 3324 3325 3326 3327 3328 3329 3330 3331 3332 3333 3334 3335 3336 3337 3338 3339 3340 3341 3342 3343 3344 3345 3346 3347 3348 3349 3350 3351 3352 3353 3354 3355 3356 3357 3358 3359 3360 3361 3362 3363 3364 3365 3366 3367 3368 3369 3370 3371 3372 3373 3374 3375 3376 3377 3378 3379 3380 3381 3382 3383 3384 3385 3386 3387 3388 3389 3390 3391 3392 3393 3394 3395 3396 3397 3398 3399 3400 3401 3402 3403 3404 3405 3406 3407 3408 3409 3410 3411 3412 3413 3414 3415 3416 3417 3418 3419 3420 3421 3422 3423 3424 3425 3426 3427 3428 34

FIGURE 150

GGCCCGAGGACTGCTATGGCTTCCTTTGTTGTTACCCCGGTCTGCGTCATGTTAAATCCAATGTCCTCTGTG
GTTAACTGCTCTTGGCCATCAAGTTCACCTTCATTGACAGCCAAGCAGATCCAGTTGTCAACACAAATTTATGG
CAAAATCCGGGGCTTAAGAACACCGTTACCCAATGAGATCTTGGGTCCAGTGAGCAGTACTTAGGGGTCCCCTA
TGCTCACCCCCACTGGAGAGAGCGGTTTCAGCCCCAGAACCCCCGTCTCTGGACTTGCATCCGAAATAC
TACTCAGTTGCTGCTGTGTGCCCCAGCACCTGGATGAGAGATCCTTACTGTCATGACATGCTGCCCATCTGGTT
TACCGCAATTTGGATACCTTTGATGACCTATGTTCAAGATCAAAATGAAGACTGCCCTTACTTAAACATCTACGT
GCCACCGGAGATGTGAGCCCAACAAAGAAAAACGAGATGATATAACAGATTAATGACCGTGGTGAAGACGAAGA
TATTCATGATCAGAACAGTAAAGAGCCCGTCATGGTCTATATCCATGGGGGATCTTATCATGGAGGGCACCGGCAA
CATGATTGACGGCAGCATTTTGGCAAGCTACGGAACCGTCATCGTATCACCATTAACTACCGTCTGGGAATACT
AGGTTTAAAGTACCGGTGACCGAGCAGCAAAAGGCAACTATGGGCTCTCGGTGATCAGATTCAAGCACTGCGGTG
GATTGAGGAGATGTGGAGCCCTTGGCGGGGACCCCAAGAGAGTGACCATCTTTGGTCTGGGGGTCTGGGCGCTC
CTGTGTCAGCCTGTGTACCCCTGTCCCACTACTCAGAAGGTCTCTTCCAGAAGGCCATCATTCAGAGCGGCACCGC
CCTGTGTCAGCTGGGCAGTGAATACCAAGCGGCAAGTACATCTCGGATATTGGCAGACAAGGTCCGCTGCAACAT
CTACCGACACCAAGCATGGTAGAATGCTGCGGAACAAGAACTACAAGGAGCTCATCCAGCAGACCATCAACCC
GGCCACCTACACATAGCCTTGGGCGGGTGTATCGACGGGAGCTCATCCAGAGCAGCCCCAGATCTGATGGA
GCAAGGCGAGTTCTCAACTACGACATCATGCTGGGCGTCAACCAAGGGGAAGGCTGAAAGTCTGTGACGGCAT
CGTGGATAACGAGGACGGTGTGACGCCCAACGACTTTGACTTCTCCGTGTCACACTCTGTGGACAACCTTTAGG
CTACCTTGAAGGGAAAGACACTTTGCGGGAGACTATCAAGTTCATGTACACAGACTGGGCGGATGAAGAAAAACC
GGAGACCGCGGGAAAAACCTGGTGGCTCTCTTACTGACCACAGTGGGTGCGGCCCGGCTGGCGCCGACCT
GCAGCGCGAGTACGGCTCCCCACCTACTCTTATGCTTCTATCATCTGCGCAAGCGAATGAAGCCAGCTG
GGCAGATTCCGCCCATGGTGTAGAGGTCCCCATGTCTTCCGCATCCCCATGTACGGTCCACCGAGCTCTTCAG
TTGTAACTTTTCAAGAACGAGCTCATGCTCAGCGCGGTGTCATGACATCTGACGAACTTGCCTTAAACATGG
TGATCCAAATCAACCAATCTCTCAGGATACCAAGTTCATTACACAAAAACCAACCGCTTTGAAGAGTGGCGTG
GTCCAGATTAATCCCAAGACAGCTCTATCTGCATATTGGCTTGAACCCAGAGTGAGAGATCACTACCGGGC
AACGAAAGTGGCTTTCTGTGGAACTCGTTCTCATTTGCACAACCTGAAACGAGATATTCAGTATGTTTCAAC
AACCACAAAGGTTCTCTCACAGACATGACATCATTTCCCTATGCGACCCGGCGATCTCCCGCCAGATATGSCC
AACCACAAAGCCGCCAGCAATCACTCTGCAACAATCCCAAACTCTAAGAGCCCTCAAAAAACAGGGCCCTGA
GGACACAATGTCTCTCAATTGAAACCAACAGAGATTATTCACCGCAATTAAAGTGTACCAATTGCGCTCGGGCGCTC
GCTCTCTTCTCAACATCTTAGCTTTTGGCGCGCTGTACTACAAAAAGGACAAGAGGGCGCCATGAGATCTCAG
GCGCCCGAGTCCCGCAGAGAAACACCACAATGATATCGCTCACAATCAGAACGAAGAGATCATGTCTCTGCAGAT
GAAGCGCTGGAACAACGATACAGAGTGTGAGTCGCTGCAGGCACACGACACATGAGGCTCACTTGCCTGCCAGA
CTACACCTCAGCTGCGCCGGTGCAGAGATGACATCCCACTTATGACGCCAAACACCATCAACATGATCCAAA
CACATGACGGGGATGCAAGCTTTGCACACTTTTAACACCTTCAGTGGAGGACAAAACAGTACAAATTTACCCCA
CGGACATTCACCACTAGAGTATAGCTTTGCCCTATTTCCTTCTATCCTCTGCGCTACCCGCTCAGCAACAT
AGAAGAGGGGAAGAAAGAGAGAAAGGAAGAGAGAGAGAGAAAGTCTCCAGACCGAAGATGTTTGTGCTCCCAT
GACTTAAAGCAAAATGCAAAAGGCAAGTATCCCATCCCGCAGACCTTATCGTTGGTGTGTTTCCAGTATTAC
AAGATCAACTTCTGACCTGTGAATGTGAGAAATGACACATTTCTGTAAATAACTGCTTTAAGATCTCTACCA
CTTCAATCAATGTTTAGTGTGATAGGACATCAACATTTCAAGGCCCGCGGTGTTTCCAAAGTCAATGGAAGCAGCT
GACACTCTGAAACTCAGCCAGGACACTTGATATTTTTTAATTAATGGAAGTTTAAACATTTCTTTCTGTGC
CACCAATGGATGGCTCTCTCTTAAGTGAAGAAAGAGTCAATGAGATTTTGGCCAGCATGGAAGCTGTAATCCAG
AGAGAAAGGAAAGTGAATAATTTATTATTAAGAAATGGACTGCGACGCAAACTGTACGGTCTGTGCAAAAGAG
GTGTTTTCGACGCTGAACATATATTAAAGAGCTTTGT

10017191.102401

FIGURE 151

MLNSNVLLWLTLALAIKFTLIDSQAQYPVVNTNYGKIRGLRTPLPNEILGPVEQYLGVPYASP
PTGERRFQPPPEPPSSWTGIRNTTQFAAVCPQHLDERSLLHMDLPIWFTANLDTLMTYVQDQN
EDCLYLNIYVPTEDGANTKKNADDITSNDRGEDEDIHDQNSKKPVMVYIHGGSYMEGTGNMI
DGSILASYGNVIVITINYRLGILGFLSTGDQAAKGNYGLLDQIQALRWIEENVGAFGGDPKR
VTIFGSGAGASCVSLLTLSHYSEGLFQKAI IQSGTALSSWAVNYQPAKYTRI LADKVCNML
DTTDMVECLRNKNYKELIQQTITPATYHIAFGPVIDGDVIPPDDQIILMEQGEFLNYDIMLGV
NQGEGLKFVDGIVDNEDGVTPNDFDFSVSNFVDNLYGYPEGKDTLRETIKFMYTDWADKENP
ETRRKTLVALFTDHWVAPAVAADLHAQYGSPTYFYAFYHHCQSEMKP SWADSAHGDEVFVYV
FGIPMIGPTELFSCNFSKNDVMSAVVMTYWTNFAKTGDPNQFPVQDTKFIHTKPNRFEEVA
WSKYNPKDQLYLHIGLKPRVRDHYRATKVAFWLELVPHLHNLNEIFQYVSTTTKVP PPDMTS
FPYGTTRRSPAKIWPTTKRPAITPANNPKHSDPHKTGPEDTTVLIETKRDYSTE LSVTI AVG
ASLLFLNILAFAALYYKKDKRRHETHRRPSPQRNTTNDIAHIQNEEIMSLQMKQLEHDHECE
SLQAHDTLRLT CPPDYTLTLRRSPDDIPLMTPNTITMIPNTLTGMQPLHTFNTFSGGQNSTN
LPHGHSTTRV

Signal sequence:

amino acids 1-24

Transmembrane domains:

amino acids 189-204, 675-692

FIGURE 152

GGGAAAGATGGCGGCGACTCTGGGACCCCTTGGGTCTGGCAGCAGTGGCGGCGATGTTTGT
CGGCTCGGGATGGGTCCAGGATGTTACTCCTTCTTTTGTGGGTCTGGGCAGGGGCCA
CAGCAAGTCCGGGCGGGTCAAACGTTTCGAGTACTTGAACCGGGAGCACTCGCTGTCTGAAGCC
CTACCAGGGTGTGGGCACAGGCACTTCTCTCACTGTGGAATCTGATGGGCAATGCCATGGTGA
TGACCCAGTATATCCGCCTTACCCAGATATGCAAAGTAAACAGGGTGCCTTGTGGAACCGG
GTGCCATGTTTCTTGAGAGACTGGGAGTTGCGAGGTGCACCTCAAATCCATGGACAAGGAAA
GAAGAATCTGCATGGGGATGGCTTGGCAATCTGGTACACAAAGGATCGGATGCAGCCAGGGC
CTGTGTTTGGAAACATGGACAAATTTGTGGGGCTGGGAGTATTTGTAGACACCTACCCCAAT
GAGGAGAAGCAGCAAGAGCGGGTATTCCTTACATCTCAGCCATGGTGAACAACGGCTCCCT
CAGCTATGATCATGAGCGGGATGGGCGGCCCTACAGAGCTGGGAGGCTGCACAGCCATTGTCC
GCAATCTTCATTACGACACCTTCTCTGGTGATTCGCTACGTCGAAGAGGCATTTGACGATAATG
ATGGATATTGATGGCAAGCATGAGTGGAGGGACTGCATTGAAGTGCCCGGAGTCCGCCTGCC
CCGCGGCTACTACTTCCGCACCTCCTCCATCACTGGGGATCTCTCAGATAATCATGATGTCA
TTCCTTTGAAGTTGTTTGAACCTGACAGTGGAGAGAACCCCAAGAGAAAGGCTCCATCGA
GATGTGTTCTTGGCCTCAGTGGACAAATATGAAGCTGCCTGAGATGACAGCTCACTGCCGC
CCTGAGTGGCCTGGCCCTCTTCTCTCATCGTCTTTTTCTCCCTGGTGTCTTCTGATTTTGCCA
TAGTCATTGGTATCATACTCTACAACAAATGGCAGGAACAGAGCCGAAAGCGCTTCTACTGA
GCCCTCTGTCTGCCACCACTTTTGTGACTGTCACCCATGAGGTATGGAAGGAGCAGGCACTG
GCCTGAGCATGCGACCTGGAGAGTGTTCTTGTCTCTAGCAGCTGTTTGGGACATATATTCTG
TCACTGGAGTTTTTGAATCAGGGACCCCGCATTTCCCATGGTTGTGATGGGGACATCTAACT
CTGGTCTGGGAAGCCACCCAGGGCAATGTGCTGTGATGTGCCCTTCCCTGCAGTCC
TTCATGTGGGAGCAGAGGTGTGAAGAGAATTTACGTGGTTGTGATGCCAAAATCAGAGAAC
AGAATTTTCATAGCCCAGGCTGCCGTGTTGTTTGACTCAGAAGGCCCTTCTACTTTCAGTTTGT
AATCCACAAGAAATTAATAAATCTGGTAACACCACAGGCTTCTGACCATTCCATTCTGGGTT
TTGCATTTGACCCAACCCCTCTGCCTACCTGAGGAGCTTCTTGTGAAACAGGATGGAACCT
TCTTCCCTGCCTTACCTTCCTTTCACCTCCATTCACTGTCTCTGTGTGCAACCTGAGCTG
GGAAGGCATTTGGATGCCCTCTGTGTGGGGCTGGGGCTGCAGAACACACCTGCGTTTCAC
TGGCCTTCATTAGGTGGCCCTAGGGAGATGGCTTTCTGCTTTGGATCACTGTTCCCTAGCAT
GGGTCTTGGGTCTATTGGCATGTCCATGGCCTTCCCAATCAAGTCTCTTCAGGCCCTCAGTG
AAGTTTGGCTAAGGTTGGGTGAAAAATCAAGAGAAAGCCTGGAAGACATCATGGATGCCATG
GATTAGCTGTGCAACTGACCACTCCAGGTTTGATCAAACCAAAGCAACATTGTGTCATGTG
GTCTGACCATGTGGAGATGTTTCTGGACTTGCTAGAGCCTGCTTAGCTGCATGTTTTGTAGT
TACGATTTTGGAAATCCACTTTGAGTGCTGAAAGTGAAGGAAGCTTCTTCTTACACCTT
GGGCTTGGATATTGCCAGAGAAGAAATTTGGCTTTTTTTTCTTAATGGACAAGAGACAGT
TGCTGTCTCTATGTTCCAAGTCTGAGAGCAACAGACCTCATCATCTGTGCTGGAGAGGTT
CACTGTCAATTGAGCAGACGCCCTGAGTGTGCCTCTGTCAACCTTATTCCACTGCTTCA
TTTGACAAGGGGTTACATGCTGCTCACCTTACTGCCCTGGGATTAAATCAGTTACAGGCCAG
AGTCTCCTTGGAGGGCTGGAACCTCTGAGTCTCTTATGAACCTCTGTAGCCTAAATGAAAT
TCTTAAAAATCACCAGTGAACCAAAAAAAGGGCGGCGGCGACTCTAGAGTCG
ACCTGCAGTAGGGATAACAGGGTAATAAGCTTGGCCGCATGG

1017191.102401

FIGURE 153

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50911

><subunit 1 of 1, 348 aa, 1 stop

><MW: 39711, pI: 8.70, NX(S/T): 1

MAATLGPLGSWQQWRRCLSDRGSRMLLLLLLLGSGQGPGQVAGQTFEYLKREHSLSKPYQ
GVGTGSSSLWNLMGNAMVMTQYIRLTPDMQSKQGALWNRVPCFLRDWELQVHFQKHGQKKN
LHGDGLAIWYTKDRMQPGPVFGNMDKFGVGLGVFVDTPNEEKQQERVFPYISAMVNNGSLSY
DHERDGRPTELGGCTAIVRNLYHDTFLVIRYVKKRHLTIMMDIDGKHEWRDCIEVPGVRLPRG
YYFGTSSITGDLSDNHDVISLKLFEFTVERTPEEEKLHRDVFPLPSVDNMKLPENTAPLPPLS
GLALFLIVFFSLVFSVFAIVIGIILYNKWQEQSRKRFY

Signal sequence:

amino acids 1-38

Transmembrane domain:

amino acids 310-329

FIGURE 154

CCGAGCCGGGCGCGCAGCGAGCGAGCTGGGGCGGGCTGGGACCATGGGCGTGAGTGCAATCTACGGATCAGTCT
 CTGATGTGGTGGTTCGTTAACTCTCAGTGGGACCTCCAAGATTTCCATGAAGAAAAATCAGTTGTCTTCATTCAAGAAAT
 TGGGGTCTGGCTCAGAATTCCTCAGCTGGTGAAAATCTGTTTTCTAGAAGAGGTTTAATTAATGCCTGCAGTCT
 GACATGTTCCCGATTTCAGGTGAAACCATGAAGAGAAAAATAGATACCTTAATAATGCTTTTCCGCAACCGCTCTCT
 TGCTGCTGCTGGCCCTGGCTGCCTGCTGGCCCTTTGTGAGCCTCAGCTGCACTCTTCCACCTGATCCCGGTGT
 CGACTCTTAAGAAATGGAATGAGTAGCAAGAGTCGAAAGAGAAATCATGCCGACCCCTGGACGGAGCCCTGTGA
 CAGACCCCGTTTATGAAGCTCTTTGTACTGCACATCCCCAGTGTGGCCGAGGCGCAGATGGAAGGTCAATGCC
 CGCATCATTTTAAAGCTGGTCTCAGTGCATGTGTTTCATGCCACGGAGACAGGTACCCACTGTATGTCATTGCCA
 AAACAAGCGACAGAAATGACTGCATCTGGTGGCTAAACAGGAAACCGTATACCCAAAATGGAAGCTTTCA
 TTAGTCACATGTCAAAAGGATCCGGAGCTCTTTTCGAAAGCCCTTGAACCTCTTTCCTTTACCCAAATCAC
 CATTTGTGTGAGATGGGAGAGCTCACACAGACAGGAGTTGTGCAGCATTTCGAGAAACGGTCAGCTGCTGAGGGATA
 TCTATCTAAAGAAACACAAATCCTGCCCCAATGATTGGTCTGCAGACCAAGCTCTATTATTAGAGACCACTGGGAAAA
 GCCGGACCTTACAAAGTGGGCTGGCCTTGTCTTTATGGCTTTCTCCAGATTTCGATGGAAGAGATTATTTCA
 GGCACCAAGCAAGTGCCTGTTCTGCTCTGGAAGCTGCTATTGCCCGTAAGAAACCGATATCTGGAAGAGGAGC
 AGCGTCGTGATACCTCTTACGTTTGAAGAACAGCCAGCTGGAGAGACCTACGGGGAGATGGCCAAGATCGTGG
 ATGTCCCAACCAAGCAGCTTAGAGCTGCCAACCCCATAGACTCCATGCTCTGCCACTCTTGCCACAATGTCACT
 TTCCCTGTACCAAGAAATGGCTGTGTTGACATGGAGCACTTCAAGGTAAATTAAGACCCATCAGATCGAGGATGA
 GGGAAAGACGGGAGAGAAATTTGACTTTCGGGTATTCTCTCCTGGGTGCCACCCCATCTCTGAACCAAAACCTCG
 GCGGGATGCAGCGTGCCACCGAGGGCAGGAAAGAGGCTCTTGGCCCTCTACTCTGCTCATGATGTCACTCTGT
 CACCAGTTCTCAGTGCTTGGGCTTTTCAGAAGCCAGGTTCCCAAGGTTTGACGCCAGGTGATCTTTGAGCTTT
 GGCAGACAGAGAAAGCCAGCTGAACATTCCGTCCGGATTCTTTACAATGGCGTGGATGTGCATCTTCCACACT
 CTTTCTGCGCAAGACCACCAAGCGTTCTCCCAAGCCATGTGCCCGCTTGAAACCTTGGTCCGCTTTGTGAAAA
 GGGACATGTTTGTAGCCCTGGGTGGCAGTGGTCAAAATATTATGATGCATGTACGACGGAAGGATTCTAAAGG
 TATGCACTACAGCAGTATGAAATCATGCCCAATACAGAGCATAGGAAAGGTCACCTCTAGTTTGTCTGTGAT
 TAGGGTAGAAGATTATTTGCTTTTAAAGCTTAATATTGTTTGTGGGAACACAGAGTGTTGGGGTTGAACACT
 AAGCAGATTGCTGCATGTGGTACGTGAATGCTTGGTACAAAATGGCCAGTTTACAGAGGAATAGAAGGTACTT
 TATCTACGCGACGACTCTGCTTAGAATGCCAAGAAATAATAGTTCAAGACCTGAAAGTGCCTTGCCTTGCAGT
 CTTTCTGCGCCAGTCCCTGTTACTATGTGATGGAACAGCACACTCAACCAAAATTTTTTTTAACTCTTAGACATT
 TTTACTCTGCTCTGTTAAAGAAATTTCTGAAAGTGAATTTCTAAATAAAGGTTGGCAAACTTTTTCTGTAAGG
 GCCAGATTGTAAATATTGAGCTGTGTGACCAAAAGGCCACATACAGTCTCTGTCAATACTACTCAACTCTGT
 TTTCTGAAGACGGAAGGCCACACAGACAGTACATAAAGGAATAATGTGTAGCTGGGTTCCAGGCCAGACAAAAA
 GATGGTGACCAAGCTTGGCCCTGGGCTGTAGTTTGTCTGACCCCTCATCTAAAAAATAGGCTATACATAAATG
 ACTTCCAGCACTTTGAGACAGTGTGAATACCAAGAAATTAATCAATGGTTCTCCAGTAACCTCTGCTAGAAACA
 CAGAAATTTGGTCTGTATCTGACACTAGAACAAAATCTGAGGGTAAATAAATCAATGAAATAGAATGAATCATAGA
 AACTGATTAGAAAGAACTCTGATGTTTATGATGATTGTGGTACAAGATGTTTTAAATGATTTCTCAATATTTGT
 TTGCTGTAGTCTATTGTGCTATATGCTGAAATTTTTGTATGCCATTTAGTATTTTTATAGTTTAGGAAAAATATT
 CTCTAAGACCACTTTTAGATGACTCTTATTCCTGTAGTAATATTCAATTTGCTGTACCTGCTGGTTGGTTAGAG
 GAGGCTAGAAGATGAATTCAGGCACCTTTCTCCAATAAACTAATTTATGGCTCATTCCTTTGACAAGCTGTAGA
 ACTGGATTCTTTTAAACCAATTTTCATCAGTTTCAAATGGTAAATTTCTGATGATTTTTAAATGCGTTTTTGA
 AGAATTTGCTATTAGGTAGTTTACAGATCTTTATAAGGTGTTTTATATTAGAAGCAATTTATATTTACATCTG
 TGATTTCTGAATAATGGTGTCTAATTCAGAGAAATGGAAGTGAAGTGATTTCTGTGTCTCCGCACTCC
 AACTTTTCTCTTTTGTGTTTGTCCAGTGTGCATTGGAATATGCTGTTTCTATAAATAAATTTTTTAGAATAA

10037471.10037471

FIGURE 155

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48329
><subunit 1 of 1, 480 aa, 1 stop
><MW: 55240, pI: 9.30, NX(S/T): 2
MLFRNRFLLLALAALLAFVSLSLQFFHLIPVSTPKNGMSSKSRKRIMDPVTEPPVTDPVY
EALLYCNIPSV AERSMEGHAPHHFKLVSVHVFIRHGDRYPLYVIPKTKRPEIDCTLVANRKP
YHPKLEAFISHMSKSGSGASFESPLNSLPLYPNHPLCEMGELTQTGVVQHLQNGQLLRDIYLK
KHKLLLPNDWSADQLYLETTGKSRTLQSGLALLYGFLPDFDWKKIYFRHQPSALFCSGSCYCP
VRNQYLEKEQRRQYLLRLKNSQLEKTYGEMAKIVDVPTKQLRAANPIDSM LCHFCHNVSFPC
TRNGCVDMEHFKVIKTHQIEDERERREKKLYFGYSLG AHPILNQ TIGRMQRATEGRKEELF
ALYSAHDVTLSPVLSALGLSEARFPRFAARLIFELWQDREK PSEHSVRILYNGVDVTFHTSF
CQDHHKRSKPKMCPLENLVR FVKRDMFVALGGSGTNYDACHREGF
```

Signal sequence:

amino acids 1-18

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 156

AAAAAAGCTCACTAAAGTTTCTATTAGAGCGAATACGGTAGATTCCATCCCCCTTTGAAGAACAAGTACTGTGGA
 GCTATTTAAGAGATAAAAAAGAAATATCCTTTCTGGGAGTCAAGATTGTGCAATTAATTGGTTAGGACTCTGAGC
 CGCCCTGTTCAACCAATCGGGAGAGAAAAGCGAGATCCTGCTCGCTTGCACGCGCTTGAGGACAAAGCAGAT
 AGCTAGGAATGAACCATCCCTGGGAGTATGTGGAACAACCGAGGAGCTCTGACTTCCCAACTGTGCCATTTCTAT
 GGGCGAAGGAAGTCTCTGACTCTCAGTGGTTAAGGGCAGAATTGAAAATAATTCTGGGAGAAGATAAGATGATGAT
 TCCTCGCGACTGCAACCGGACTACAAGGGCTGTGCTGCTGGGAATCCTCTGGGGACTCTGTGGGAGACCGG
 ATGCAACCCAGATAACGCTATTAGTTCGGGAAGAGCTGGGAAAAGGCTCTAGGGTGGGCGACATCTCCAGGGACCT
 GGGGCTGGAGCCCGGGAGCTCGCGGAGCGCGGAGTCCGATCATCCCAAGGTTAGGACGAGCTTTTCGCCCT
 GAATCCGCGCAGCGGAGCTTGGTCAACGGCGGCGAGGATAGACCGGGAGGAGCTCTGTATGGGGCCATCAAGTG
 TCAATTAAATCTAGACATTTCTGATGGAGGATAAAGTGAAAATATATGGAGTAGAAGTAGAAGTAGAGGACATTAA
 CGACAATGCGCCTTACTTTCTGTAAGAGTGAATTAGAAAATAAAATAGTGAATAATGCAGCCATGAGATCGGTT
 CCCTCTACCCCAACGCTGGGATCCGGATATCGGGAAGAATCTCTGCAGAGCTACGAGCTCAGCCGGAACACTCA
 CTTCCTCCCTCATCGTGAATAATGGAGCGACGGTAGTAAGTACCCGGAATTGGTGTGTAAGACGCGCCCTGGAGCG
 CGAAGAAAAGGCTGCTCAACACCTGGTCTTACGGCTCCGACGGGGCGACCCGGTGCACAGGCAACCGCGCG
 CATCCGCGTGATGGTTCTGATGCGAAGCAACACGACCCAGCGTTTGTCTCAGCCCGAGTACCGCGCGAGCGTTCC
 GGAGAATCTGGCCTTGGGCACGCGAGCTGCTGTAGTCAACGCTACCGACCTGACGAAGGAGTCAATGCGGAAGT
 GAGGTATTCCTCCGGTATGTGGAGCAACAGGCGGCCAAGTTTCAAACCTAGATTTGAATTCAGGGAACAATATC
 AACAAATAGGGGAGTTGGACCAAGAGGAGTCAAGATTCTACCAAGTGGAACTGAAGCAATGGATAATGCAAGGATA
 TTCTGCGCGAGCCAAAGTCTGATCACTGTTCTGGACGTGAACGACAATGCCCAGAAGTGGTCTCCTCACTCTCT
 CGCCAGCTCGGTTCCCGAAAATCTCTCCAGAGGGACATTAATTGCCCTTTAAATGTAAATGACCAAGATTCTGA
 GGAAAACGAGCAGGTGATCTGTTCATCCCAAGGAAATCTGCCCTTTAAATAGAAAATCTTACGGAATTTACTA
 TAGTTTAGTACAGACATAGTCTTGGATAGGGAACAGGTTCTAGCTACAACATCACAGTGAACCGCCACTGACGG
 GGGAAACCCGCCCCATCCACGGAATCTCATATCTCGCTGAACGTGGCAGACACCAACGACAACCCGCGCTATC
 GTCTACGTGTCCATCACTCCGACACTGGGGTACTGTATGGCTGAGCTCTCTCGACTCAGAGCAGTTTCCGAGT
 CTTGCAAGTGAAGTGAATGGCGCGGACAACGGGCAACCCGCGCTCCGACGAGCAGTTCGTTGAGCGTCTTGTG
 GCTGACCAAGAACAGCATGTGCGCCGAGATCCTGTACCCGCGCTCCCGACGAGCGGTTCCACTGGCGTGGAGCT
 GGTCTCCCCGCTCCGACAGCCCGGCTACTGTTGACCAAGTGGTGGCGGTGGACAGAGACTCCGCGCAGAAGCT
 CTGGCTGTCTTACCGTCTGCTCAAGCGCACGCGAGCCGGGACTCTTCTGGTGGGCTGTCACACGGGCGAGGTGCG
 CACGGCGAGAGCCCTCTGGACAGAGCGCGCTCAAGCAGAGGCTCGTAGTGGCGTCCAGGACACCGGCGAGCC
 CCCTCTCTCCGCACTGTCAACGCTCAACGCTGGCGTGGCGGACAGCATCCCCAAGTCTGGCGGACCTCGGCGAG
 CCTCGAGTCTCCAGCTAACTTGAAACCTCAGACCTCACTCTGTAACCTGGTGGTAGCGGTGGCGCGGACTCTCTG
 CGTCTCTCTGGCCTTCTCATCTTCTGCTGCGGCTCAGGCTCGCGCGCTCGGCAACAGTCAACGCTCTGTGAGAG
 TTAGAGAGCGCGCTTGACAGAGGCGCGCGCTCAGCTTTGAGGAGTCAACCTGATCTTCCCCAGCCCACTATGCAAG
 CTATTCCCAACAGGTTTCCCTCACCAGGACTCGCGGAAGAGTCACTGATCTTCCCCAGCCCACTATGCAAG
 CAGTCTGTTGAATGAGGTGAGTTTATATAAATCTCTTCTTTTCTTTTAAATGCTCTGTCTCCCAAGC
 TGGAGTGCAGCGGTACGATCATAGCTCACTGCGGCTCAAACTCCTAGGCTCAAGCAATTATCCCACTTTGCT
 CCGGTGTAAACAGGACTACAGGTCAAGCCACTACTGTCTGCTATCTATCTATCTATCTATCTATCTATCTAT
 CTATCTATCTATCTATCTATCTATCTTCTGTACAGACGGGAGTCTACGCGCTTAATCCAGTACTTTGGGAGCG
 CGAGGCGGGTGAATCACTGAGGTTGGAGTTTGAGACCAAGCTGACCAACATGGAGAAAACCGGCTCTATACTAA
 AAAAAATCAAAATAGCGCGGTGGTGGTGCATGCTGTAAATCCAGCTACTTGGGAGGCTGAGTCAGGAGAT
 TGCTTTAACCTGGGAGGTGGAGGTTGCAATGAGCTGAGATTGTGCCATTGCACCTCAGCCTGGGCAACAAGAGTG
 AAATCTATCTCA

10037771.100401

FIGURE 157

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48306
><subunit 1 of 1, 916 aa, 1 stop
><MW: 100204, pI: 4.92, NX(S/T): 4
MIPARLHRDYKGLVLLGILLGLWETGCTQIRYSVPPELEKGSRVGDISRDGLGLEPRELAER
GVRIIPRGRTQLFALNPRSGSLVTAGRIDREELCMGAIKCQLNLDILMEDKVKIYGVVEVEVR
DINDNAPYFRESELEIKISENAATEMRFPPLPHAWDPDIGKNSLQSYELSPNTHFSLIVQNGA
DGSKYPELVLKRALDREEKAAHHLVLTASDGGDPVRTGTARIRVMVLDANDNAPAFAPQPEYR
ASVPENLALGTQLLVNATDPDEGVNAEVRYSFYVDDKAAQVFKLDCNSGTISTIGELDHE
ESGFYQMEVQAMDNAGYSARAKVLI'VLDVNDNAPEVVLTSLASSVPENSPRGTLIALLNVN
DQDSEENGQVICFIQGNLPFKLEKSYGNYYSLVTDIVLDREQVPSYNITVTATDRGTPPLST
ETHISLNVADTNDNPPVFPQASYSAYIPENNPRGVSLSVTAHDPPDCEENAQITYSLAENTI
QGASLSSYSVINSDTGVLYALSSFYEQFRDLQVKVMARDNGHPPLSSNVSLSLFVLDQNDN
APEILYPALPTDGSTGVELAPRSAEPGYLVTKVVAVD RDSQNAWLSYRLLKASEPGLFSVG
LHTGEVTRARALLDRDALKQSLVVAVQDHGQPPLSATVTLTVAVADSIQVLADLGSLESPA
NSETSDLTLYLVVAVAAVSCVFLAFVILLALLRLRRWHKSRLQASGGGLTGAPASHFVGVD
GVQAFQLQTSHEVSLTTDSRKSHLIFPQPNYADMLVVSQESFEKSEPLLLSGDSVFSKDSHGL
IEVSLYQIFFLFFNCSSVSQAGVQRYDHSSLRPQTPRLKQLSHLCLRCNDRDYCKPPTVCLSI
IYLSIYLSIYLSIYLLLSCTDGS LTPVIVLWEAEAGGSPEVGSLRPA
```

Signal sequence:

amino acids 1-30

Transmembrane domains:

amino acids 693-711, 809-823, 869-888

FIGURE 158

CCCAGGCTCTAGTGCAGGAGGAGAAGGAGGAGGAGCAGGAGGTGGAGATTCCCAGTTAAAAG
GCTCCAGAATCGTGTAACAGGCAGAGAACTGAAGTACTGGGGCCTCCTCCACTGGGTCCGAA
TCAGTAGGTGACCCCGCCCTGGATTCTGGAAGACCTCACCAATGGGACGCCCCGACCTCGT
GCGGCCAAGACGTGGATGTTCTGTCTCTTGCTGGGGGAGCCTGGGCAGGACACTCCAGGGC
ACAGGAGGACAAGGTGCTGGGGGGTCATGAGTGCCAACCCATTTCGAGCCTTGGCAGGCGG
CCTTGTTCCAGGGCCAGCAACTACTCTGTGGCGGTGTCTTGTAGGTGGCAACTGGGTCCTT
ACAGCTGCCCACTGTAAAAAACCGAAATACACAGTACGCCTGGGAGACCACAGCCTACAGAA
TAAAGATGGCCAGAGCAAGAAATACCTGTGGTTCAGTCCATCCCACACCCCTGCTACAACA
GCAGCGATGTGGAGGACCACAACCATGATCTGATGCTTCTTCAACTGCGTGACCAGGCATCC
CTGGGGTCCAAAGTGAAGCCCATCAGCCTGGCAGATCATTGCACCCAGCCTGGCCAGAAGTG
CACCGTCTCAGGCTGGGGCACTGTCAACAGTCCCCGAGAGAAATTTTCTGACACTCTCAACT
GTGCAGAAGTAAAAATCTTTCCCAAGAAGTGTGAGGATGCTTACCCGGGCAGATCACA
GATGGCATGGTCTGTGCAGGCAGCAGCAAAGGGGCTGACAGTGCCAGGGCGATTCTGGAGG
CCCCCTGGTGTGTGATGGTGCCTCCAGGGCATCACATCCTGGGGCTCAGACCCCTGTGGGA
GGTCCGACAAACCTGGCGTCTATACCAACATCTGCCGCTACCTGGACTGGATCAAGAAGATC
ATAGGCAGCAAGGGCTGATTCTAGGATAAGCACTAGATCTCCCTTAATAAACTCACAACTCT
CTGGTTC

FIGURE 159

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48336

<subunit 1 of 1, 260 aa, 1 stop

<MW: 28048, pI: 7.87, NX(S/T): 1

MGRPRPRAAKTWMFLLLLGGAWAGHSRAQEDKVLGGHECQPHSQPWQAALFQGGQQLLCGGVL
VGGNWWLTAAHCKKPKYTVRLGDHSLQNKDGPEQEIPIVVSQSIHPHCYNSSDVEDHNHDLMLL
QLRDQASLGSKVKPIISLADHCTQPGQKCTVSGWGTVTSPRENFPDTLNCAEVKIFPQKKCED
AYPGQITDGMVCAGSSKGGADTCQGDSSGGPLVCDGALQGITSWGSDFCGRSDKPGVYTNICRY
LDWIKKIIGSKG

Important Features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 51-71

N-glycosylation site.

amino acids 110-113

Serine proteases, trypsin family, histidine active site.

amino acids 69-74 and 207-217

Tyrosine kinase phosphorylation site.

amino acids 182-188

Kringle domain proteins motif

amino acids 205-217

FIGURE 160

100
90
80
70
60
50
40
30
20
10
0

GGCGCCGCTGCACCGGGCGGGCTGAGCGCCTCCTGCGGCCCGGCCCTGCGCGCCCCGGGCCCGC
CGCGCCGCCACGCCCCAACCCCGGCCCGCGCCCCCTAGCCCCCGCCGGGCCCGCGCCCCG
GCCCGCGCCAGGTGAGCGCTCCGCCCGCGCGAGGCCCGCCCGCCCGCCCGCCCGCCCGC
CCCCGGCCGGCGGGGGAACCGGGCGGATTCTCGCGCTCAAACCACTGATCCCATAAAAC
ATTATCTCTCCCGCGCGCCCGCTGCGAGCGCCCCGCCAGTCCGCGCGCGCCCGCCCTCG
CCCTGTGCGCCCTGCGCGCCCTGCGCACCCGCGGCCCGAGCCAGCCAGAGCCGGGCGGAGC
GGAGCGCGCCGAGCCTCGTCCCGCGGCCGGGCCGGGGCCGTCGAGCGCGCGCTGGA
TGCGAGCCCGCGCGGGGAGACGGGCGCCGCCCGAAACGACTTTTCAGTCCCCGACGCGC
CCCGCCAAACCTACGATGAAGAGGGGCTCCGCTGGAGGGAGCCGGCTGCTGGCATGGGTG
CTGTGGCTGCAGGCTTGGCAGGTGGCAGCCCCATGCCAGGTGCCTGCGTATGCTACAATGA
GCCCCAAGGTGACGACAAGCTGCCCCCAGCAGGGCTGCGAGGTGTGCCGTGGGCATCCCTG
CTGCCAGCCAGCGCATCTTCTGCAACGCAACCGCATCTCGCATGTGCCAGCTGCCAGCTTC
CGTGCCTGCCGCAACCTCACCATCTGTGGCTGCACTCGAATGTGCTGGCCCCGAATTGATGC
GGCTGCCTTCACTGGCCTGGCCCTCCTGGAGCAGCTGGACCTCAGCGATAATGCACAGCTCC
GGTCTGTGGACCTGCCACATTCACGGCCTGGGCCGCTTACACACGCTGCACCTGGACCGC
TGCGCCTGCAAGAGCTGGGCCCGGGGCTGTTCCGCGCCTGGCTGCCTGCAGTACCTCTA
CCTGCAGGACAACGCGCTGCAGGCACTGCCTGATGACACCTTCCGCGACCTTGGGCAACCTCA
CACACCTCTTCTGCAACGGCAACCGCATCTCCAGCGTGGCCGAGCGCCCTTCCGTGGGCTG
CACAGCCTCGACCGCTCTCTACTGCACAGAACCGCGTGGCCCATGTGCACCCGCATGCCCT
CCGTGACCTTGGCCGCTCATGACACTTATCTGTTTGCACCAATCTATCAGCGCTGCCCA
CTGAGGCCCTGGCCCCCTGCGTGCCCTGCAGTACCTGAGGCTCAACGACAACCCCTGGGTG
TGTGACTGCCGGGCACGCCCACTCTGGGCTGGCTGCAGAGTTCCGCGGCTCCTCTCCGA
GGTGCCCTGCAGCCTCCCGCAACGCTGGCTGGCGGTGACCTCAAACGCTTAGCTGCCAATG
ACCTGCAGGGCTGCGCTGTGGCCACCGGCCCTTACCATCCCATCTGGACCGCGAGGCCCC
GATGAGGAGCCGCTGGGGCTTCCCAAGTGCTGCCAGCCAGATGCCGCTGACAAGGCTCAGT
ACTGGAGCCTGGAAGACCAGCTTCCGAGGCAATGCGTGAAGGGAACGCTGCCGCCCGGTG
ACAGCCCGCCGGGCAACGGCTCTGGCCCAACGGCAATCAATGACTCACCCCTTGGGACTCTG
CCTGGCTCTGCTGAGCCCCGCTCACTGCAGTGCGGCCCGAGGCTCCGAGCCACCAAGGTT
CCCCACCTCGGGCCTCGCCGGAGGCCAGGCTGTTACGCAAGAACCGCACCCGCGGCCACT
GCGCTCTGGGCCAGGCAGGCAGCGGGGCTGGCGGGACTGGTGAATCAGAAGGCTCAGGTGCC
CTACCCAGCCTCACCTGCAGCCTCACCCCCCTGGGCTGGCGCTGGTCTGTGGACAGTGT
TGGGCCCTGCTGACCCCCAGCGACACAAGAGCGTGTGCTCAGCAGCCAGGTGTGTGTACATAC
GGGGTCTCTCTCCACGCGGCAAGCCAGCCAGCGGGCGGCGGACCCGTCGGGGCAGGCCAGGCAG
GTCCTCCCTGATGGACGCTGCGCCCGGCCACCCCATCTCCACCCCATCATGTTTACAGGG
TTCCGCGGCGAGCGTTTGTTCAGAACCGCGCTCCACCCAGATCGCGGTATATAGAGATAT
GCATTTTATTTTACTTGTGTAAAAATATCGGACGACGTGGAATAAGAGCTCTTTTCTTAA
AAAA

FIGURE 161

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44184

><subunit 1 of 1, 473 aa, 1 stop

><MW: 50708, pI: 9.28, NX(S/T): 6

MKRASAGGSRLRLAWVLWLQAWQVAAAPCPGACVCYNEPKVTTSCPQQGLQAVPVGIPAAASQRI
FLHGNRISHVPAASFACRNLTIILWLSNVLARIDAAAFGLALLEQLDLSDNAQLRSVDPA
TFHGLGRHLTHLDRCLGELGPGGLFRGLAALQYLYLQDNALQALPDDTFRDLGNLTHLFLH
GNRISSVPERAFRGLHSLDRLLHQNRAHVHHPHAFRDLGRIMTYLFFANNLSALPTEALAP
LRALQYLRRLNDNPWVCDRCRARPLWAWLQKFRGSSSEVPKSLPQRLAGRLKRLAANDLQGCA
VATGPYHPITWGRATDEPLGLPKCCQPDADKASVLEPGRPASAGNALKGRVPPGDSPPGN
GSGPRHINDSPFGTLPGSAEPPLTAVRPEGSEPPGFPTSGPRRRPGCSRKNRTRSHCRLLGQA
GSGGGGTGDSESGGALPSLTCSLTPLGLALVLWTVLGPC

Important features:

Signal peptide:

amino acids 1-26

Leucine zipper pattern.

amino acids 135-156

Glycosaminoglycan attachment site.

amino acids 436-439

N-glycosylation site.

amino acids 82-85, 179-183, 237-240, 372-375 and 423-426

VWFC domain

amino acids 411-425

FIGURE 162

GGAAGTCCACGGGGAGCTTGGATGCCAAAGGGAGGACGGCTGGGTCTCTGAGAGGAGCTAC
TCACTGGCATATTTCTGAGGTATCTGTAGAATAACCAAGCCCTCAGATACCTGGGGACTTTAC
AGTCCCACAGAACCGTCTCTCCAGGAAGCTGAAATCCAGCAAGAACATGGAGGGCCAGCGGGA
AGCTCATTTGCAGACAAAGGCAAGTCTCTTTTCTCTCTTTTGGGCTTATCTCTGGCG
GGCGCGGCGGAACCTAGAAGCTATTCTGTGGTGGAGGAACTGAGGGCAGCTCTTTTGTCTAC
CAATTTAGCAAAGGACCTGGGTCTGGAGCAGAGGGAATTTCTCCAGGCGGGGGTTTAGGGTTG
TTTCCAGAGGGGAACAAACTACATTTGCAGCTCAATCAGGAGACCGCGGATTTGTGCTAAAT
GAGAAATTTGAGCCGTGAGGATCTGTGCGGTACACAGAGCCCTGTGTGCTACGTTTCCAAGT
GTTGCTAGAGAGTCCCTTCGAGTTTTTCAAGCTGAGCTGCAAGTAATAGACATAAACGACC
ACTCTCCAGTATTTCTGGACAAACAAATGTTGGTGAAAGTATCAGAGAGCAGTCTCTCTGGG
ACTACGTTTCTCTGAAGAATGCCGAAGACTTAGATGTAGGCCAAAACAATATTGAGAACTA
TATAATCAGCCCAACTCTTATTTTCGGTCTCTACCCGCAACGAGTGATGGCAGGAAAT
ACCCAGAGCTGGTGTCTGGACAAAGCGCTGGACCGAGAGGAAGAAGCTGAGCTCAGGTTAAAC
CTCACAGCAGCTGGATGTGGTCTCTCCGCCAGATCTGGCAGCTCTCAGGTCTACATCGAAGT
CTCGGATGCTCAACGATTAATGCCCTGAAATTTGAGCAGCCTTCTCTAGAGTGCAGATCTCTG
AGGACAGTCCGGTAGGCTTCTCGGTTGTGAAGGTCTCTGCCACGGATGTAGACACAGGAGTC
AACCGAGAGATTTCCTATTTCACTTTTCCAAGCTTCAGAAGAGATTGGCAAAACCTTTAAGAT
CAATCCCTTGACAGGAGAAATTGAACATAAAAAACAACCTCGATTTGAAAAAATCTTCAGTCTCT
ATGAAGTCAATATTGAGGCAGAGATGCTGGAACCTTTCTGGAATAATGCACCGTTCTGATT
CAAGTGATAGATGTGAACGACCATGCCCTCAGAGTTACCATGTCTGCATTTACCAGCCCAAT
ACCTGAGAACCGCCCTGAAACTGTGGTTGCACTTTTCAGTGTTCAGATCTTGATTACGGAG
AAAAATGGGAAAAATTAGTTGCTCCATTCAGGAGGATCTACCCCTTCTCTGAAATCCGCGGAA
AACTTTTACACCTACTAACGGAGAGACCCTAGACAGAGAAAGCAGAGCGGAATACAACAT
CACTATCACTGTCACTGACTTGGGGACCCCTATGCTGATAACACAGCTCAATATGACCGTGC
TGATCGCGGATGTCAATGAACCGCTCCCGCCTTCACCCAAACCTCTTACACCTGTTCGTC
CGCGAGAACCAACAGCCCCGCCCTGCACATCCGCGAGCGTCAGCGCTACAGACAGAGACTCAGG
CACCAACGCCACAGGTCACTTACTCGCTGCTGCCGCCCCAGGACCCGCACTGCCCTCACAT
CCCTGGTCTCCATCAACGCGGACAACCGCCACCTGTTTCGCCCTCAGGTCTCTGGACTACGAG
GCCCTGCAGGGGGTTCAGTTCCGCGTGGGCGCTTCAGACCACGGCTCCCCGGCGCTGAGCAG
CGAGGCGCTGGTGCGCGTGGTGGTCTGGACGCCAACGACAACCTCGCCCTCTGTGCTGTACCC
CGCTGCAGAACCGCTCCGCGCCCTGCACCGAGCTGGTGCCCGGGCGGCCAGCGGGCTAC
CTGGTGACCAAGGTGTGGCGGTGAGCGGCGACTCGGGCCAGAACCCCTGGCTGTCTGTACCA
GCTGCTCAAGGCCACCGAGCTCGGTCTGTTTCGGCGTGTGGGCGCAATGGCGAGGTGCGCA
CCGCCAGGCTGCTGAGCGAGCGCAGCGGCCAAGCACAGGCTGGTGGTGTGCTGGTCAAGGAC
AATGGCGAGCCTCCGCGCTCGGCCACCGCCACGCTGCACGTGCTCTGTGGAGCGGCTCTC
CAGCCCCTACTCGCTCTCCCGGAGGCGGCCGACCCAGGCCAGGCCGACTTGCTCACCG
TCTACCTGGTGGTGGCGTTGGCCTCGGTGTCTTCGCTCTTCTCTTTTCGGTGTCTCTGTTT
GTGGCGGTGCGGCTGTGTAGGAGGAGCAGGGCGGCCCTCGGTGGGTGCTGCTTGGTGCCCGA
GGGCCCCCTTCAGGGCATCTTTGTGGACATGAGCGGCACAGGACCCCTACCCAGAGCTACC
AGTATGAGGTGTGTCTGGCAGGAGGCTCAGGGACCAATGAGTTCAAGTTCCGTAAGCCGATT
ATCCCCAATCTCCCTCCCACTGCCCTGGGAAAGAAATACAAGGAAATTTACCTTCCCCAA
TAACTTTGGGTTCATATTCAGTGAACCATAGTTGACTTTTACATTCATAGGTATTTTATTT
TGTGGCATTTCCATGCCAATGTTTATTTCCCCCAATTTGTGTGTATGTAATATTGTACGGAT
TTACTCTTGATTTTCTCATGTTCTTTCTCCCTTTGTTTTAAAGTGAACTTTACCTTTATT
CCTGGTTCTT

1007191.102401

FIGURE 163

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48314

<subunit 1 of 1, 798 aa, 1 stop

<MW: 87552, pI: 4.84, NX(S/T): 5

MEASGKILICRQRQVLFSLLLGLSLAGAAEPRSYSVVEETEGSSFVTNLAKDLGLEQREFSR
RGVRVVS RGNKLHLQLNQETADLLNEKLDREDLCGHTEPCVLRQVLLSPFEFFQAE LQV
IDINDHSPVFLDKQMLVKVSESSPPGTTFFPLKNAEDLDVGQNNIENYIIISPN SYFRVLTRKR
SDGRKYPELVLDKALDREEEAELRLTLTALDGGSPPRSGTAQVYIEVLVDNDNAPEFEQPFY
RVQISEDSPVGFLVVKVSATDVDTGVNGEISYSLFQASEEIGKTFKINPLTGEIELKKQLDF
EKLQSYEVNIEARDAGTFSGKCTVLIQVIDVNDHAPEVTMSAFTSPIPENAPETVVALLFSVS
DLDSGENGKISCSIQEDLPFLLKSAENFYTLTTERPLDRESRAEYNITITVTDLGTPMLITQ
LNMTVLIADVNDNAPAFQTQTSYTLFVRENNSPALHIRSVSATDRDSGTNAQVTYSLLPPQDP
HLPLTSLVSINADNGHLFALRSLDYEALQGFQFRVGASDHGSPALSS EALVRVVLDANDNS
PFVLYPLQNGSAPCTELVPRAAEPGYLVTKVVAVDGDSGQNAWLSYQLLKATELG LFGVWAH
NGEVRTARLLSERDAAKHRLVVLVKDNGEPPRSATATLHVLLVDGFSQPYLPLPEAAPTQAQ
ADLLTVYLVVALASVSSFLFSVLLFVAVRLCRRSRAASVGRCLVPEGP LPHGLVDMSGTRT
LSQSYQYEVCLAGGSGTNEFKFLKPIIPNFPQCPGKEIQGNSTFPNNFNGFNIQ

Important features:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 685-712

Cadherins extracellular repeated domain signature.

amino acids 122-132, 231-241, 336-346, 439-449 and 549-559

ATP/GTP-binding site motif A (P-loop).

amino acids 285-292

N-glycosylation site.

amino acids 418-421, 436-439, 567-570 and 786-789

FIGURE 164

ACCCACGCGTCCGCCACGCGTCCGCCACGCGTCCGCCACGCGTCCGCGCGTAGCCGTGC
GCCGATTGCCCTCTCGGCCTGGGCAATGGTCCCGGCTGCCGCTCAGCAGCCGCCCGCGTCAT
GCGGCTCCTCGGCTGGTGGCAAGTATTTGCTGTGGGTGCTGGGACTTCCCGTCCGCGGCGTGG
AGGTTGCAGAGGAAGTGGTGGCTTATGGTCAGAGGAGCAGCCTGCTCACCTCTCCAGGTG
GGGGCTGTGTACCTGGGTGAGGAGGAGCTCCTGCATGACCCGATGGGCCAGGACAGGGCAGC
AGAAGAGGCCAATGCGGTGCTGGGGCTGGACACCCAAAGGCGATCACATGGTGATGCTGTCTG
TGATTCTCGGGGAAGCTGAGGACAAAGTGAGTTCAGAGCCTAGCGGCGTCACCTGTGGTGTCT
GGAGGAGCGGAGGACTCAAGGTGCAACGTCCGAGAGAGCCTTTTCTCTGGATGGCGCTGG
AGCACACTTCCCTGACAGAGAAGAGGAGTATTACACAGAGCCAGAAGTGGCGGAATCTGACG
CAGCCCCGACAGAGGACTCCAATAACACTGAAAGTCTGAAATCCCCAAAGGTGAACGTGTGAG
GAGAGAAACATTACAGGATTAGAAAATTTCACTCTGAAAATTTTAAATATGTCACAGGACCT
TATGGATTTTCTGAACCCAAACGGTAGTGACTGTACTCTAGTCCTGTTTTACACCCCGTGGT
GCCGCTTTTCTGCCAGTTTGGCCCCCTCACTTTAACTCTCTGCCCCGGGCATTTCCAGCTCTT
CACTTTTTTGGCACTGGATGCATCTCAGCACAGCAGCCTTTCTACCAGGTTTGGCACCGTAGC
TGTTCTTAATATTTTATTATTTCAAGGAGCTAAACCAATGGCCAGATTTAATCATACAGATC
GAACACTGGAAACACTGAAAACTTTCATTTTTTAATCAGACAGGTATAGAAGCCAAGAAGAT
GTGGTGGTAACTCAAGCCGACCAATAGGCCCTCTTCCAGCACTTTGATAAAAAGTGTGGA
CTGGTTGCTTGTATTTTCTTATCTTTTAAATAGTTTTATTATGTATGCTACCATTGAA
CTGAGAGTATTGCGTGGCTAATCCAGGACAAGAGCAGGAACATGTGGAGTAGTGATGGTCT
GAAAGAAGTTGGAAAGAGGAACCTCAATCCTTCGTTTCAGAAATTAGTGCTACAGTTTCATA
CATTTTCTCCAGTGACGTGTGTGACTTGAACTTCAGGCAGATTTAAAGAATCATTTGTTGAA
CAACTGAATGTATAAAAAAATTATAAACTGGTGTTTAACTAGTATTGCAATAAGCAAAATGC
AAAAATATTCAATAG

FIGURE 165

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48333

><subunit 1 of 1, 360 aa, 1 stop

><MW: 39885, pI: 4.79, NX(S/T): 7

MVPAAGRPRPRVMRLLGWWQVLLWVLGSLPVRGVEVAEESGRLWSEEQPAHPLQVGAVYLGE
ELLHDFPMGQDRAAEEANAVLGLDTQGDHVMVLSVIPGEAEDKVSSEPSGVTCGAGGAEDSRC
NVRESLFLSDGAGAHFPDREEEYYTEPEVAESDAAPTEDSNNTESLKSPKVNCEERNITGLE
NFTLKILNMSGDLMDFLNPNNGSDCTLVLFYTPWCRFSASLAPHFNSLPRAFPALHFLALDAS
QHSSLSTRFGTVAVPNILLFQGAKPMARFNHTDRLTLETLEKIFIFNQTGIEAKKNVVVTQADQ
IGPLPSTLIKSVDWLLVFSLSLFLISFIMYATIRTESIRWLIPGQEQUEHVE

Important features:

Signal peptide:

amino acids 1-25

Transmembrane domain:

amino acids 321-340

Homologous region to dilsufide isomerase

amino acids 212-302

N-glycosylation site.

amino acids 165-168, 181-184, 187-190, 194-197, 206-209, 278-281
and 293-296

Thioredoxin domain

amino acids 211-227

FIGURE 166

CCCGGCTCCGCTCCCTCTGCCCCCTCGGGGTGCGCGCCCCACGATGCTGCAGGGCCCTGGCT
CGCTGCTGCTGCTCTTCTCGCCTCGCACTGCTGCCCTGGGCTCGGCGCGGGGCTCTTCCCTC
TTTGGCCAGCCGACTTCTCTACAAGCGCAGCAATTGCAAGCCCATCCCGGTCAACCTGCA
GCTGTGCCACGGCATCGAATACCAGAACATGCGGCTGCCCAACCTGCTGGGCCACGAGACCA
TGAAGGAGGTGCTGGAGCAGGCCGGCGCTTGGATCCCGCTGGTCATGAAGCAGTGCCACCCG
GACACCAAGAAGTTCCTGTGCTCGCTCTTCGCCCCCGTCTGCCTCGATGACCTAGACGAGAC
CATCCAGCCATGCCACTCGCTCTGCGTGCAGGTGAAGGACCGCTGCGCCCGGTCA GTGCCG
CCTTCGGCTTCCCTGGCCCCGACATGCTTGAGTGCGACCGTTTCCCCAGGACAACGACCTT
TGCATCCCCCTCGCTAGCAGCGACCACTCCTGCCAGCCACCGAGGAAGCTCCAAAGGTATG
TGAAGCCTGCAAAAAATAAAAAATGATGATGACAACGACATAATGGAAACGCTTTGTAAAAATG
ATTTTGCACTGAAAAATAAAGTGAAGGAGATAACCTACATCAACCAGATACCAAAATCATC
CTGGAGACCAAGAGCAAGACCATTTACAAGCTGAACGGTGTGTCCGAAAGGGACCTGAAGAA
ATCGTGCTGTGGCTCAAAGACAGCTTGCAAGTGCACCTGTGAGGAGATGAACGACATCAACG
CGCCCTATCTGGTCATGGGACAGAAAAGGGTGGGGAGCTGGTGATCACCTCGGTGAAGCGG
TGGCAGAAGGGGCGAGAGAGATTCAAGCGCATCTCCCGCAGCATCCGCAAGCTGCAGTGCTTA
GTCCCCGGCATCTGATGGCTCCGACAGGCCTGCTCCAGAGCACGGCTGACCATTTCTGCTCC
GGGATCTCAGCTCCCGTTCCCAAGCACACTCCTAGCTGCTCCAGTCTCAGCTGGGCAGCT
TCCCCCTGCCTTTTGCACGTTTGCATCCCCAGCATTTCTGAGTTATAAGGCCACAGGAGTG
GATAGCTGTTTTACCTAAAGGAAAAGCCCACCCGAATCTTGTAGAAATATTCAAACATAATA
AAATCATGAATATTTTAA

FIGURE 167

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50920

><subunit 1 of 1, 295 aa, 1 stop

><MW: 33518, pI: 7.74, NX(S/T): 0

MLQGGPSLLLLFLASHCCLGSARGLFLFGQPDFS YKRSNCKPIPVNLQLCHGIEYQNMRLPN
LLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCSLFAPVCLDDLDETIQPCHS LCVQVKDR
CAPVMSAFGFPPWPDMLECDRFPQDNDLCIPLASSDHLLPATEEAPKVCEACKNKNDNDNDIM
ETLCKNDFALKIKVKEITYINRDTKIILET KSKTIYKLVGVSERDLKKSVLWLKDSLQCTCE
EMNDINAPYLVMGQKQG GELVITSVKRWQKQREFKRISRSIRKLQC

Important features:

Signal peptide:

amino acids 1-20

Cysteine rich domain, homologous to frizzled N terminus

amino acids 6-153

FIGURE 168

GTGGAGGCCGCCGACGATGCGCGGGCCGACGGAGGCCGAGACGGGGTTGGCCGAGCCCGGG
CCCTGTGCGCGCAGCGGGGCCACCGCACCTACGCGCGCCGCTGGGTGTTCTGTCTCGCGATC
AGCCTGCTCAACTGCTCCAACGCCACGCTGTGGCTCAGCTTTGCACCTGTGGCTGACGT CAT
TGCTGAGGACTTGGTCCCTGTCCATGGAGCAGATCAACTGGCTGTCTACTGGTCTACCTCGTGG
TATCCACCCCATTTGGCGTGGCGGCCATCTGGATCCTGGACTCCGTCGGGCTCCGTGCGGCG
ACCATCCTGGGTGCGTGGCTGAACCTTTGCCGGGAGTGTGCTACGCATGGTGGCCCTGCATGGT
TGTGGGACCCAAAACCCATTTGCCTTCCTCATGGGTGGCCAGAGCCTCTGTGCCCTTGCCC
AGAGCCTGGT CATCTTCTCTCCAGCCAAGCTGGTGCCTTGTGGTTCCCAGAGCACCAGCGA
GCCACGGCCAACATGCTCGCCACCATGTGCAACCTCTGGGCGTCTTGTGGCCAATGTGCT
GTCCCTGTGCTGGTCAAGAAGGTGAGGACATTCGGTTAATGCTCGGTGTCTATAACCATCC
CTGCTGGCGTCTGCTGTGCTGTCCACCATCTGCCTGTGGGAGAGTGTGCCCCCACCCCG
CCCTCTGCGGGGCTGCGAGCTCCACCTCAGAGAAAGTTCTTGGATGGGCTCAAGCTGCAGCT
CATGTGGAACAAGGCCTATGTCATCCTGGCTGTGTGCTTGGGGGAATGATCGGGATCTCTG
CCAGCTTCTCAGCCCTCCTGGAGCAGATCCTCTGTGCAAGCGGCCACTCCAGTGGGTTTTCC
GGCCTCTGTGGCGCTCTCTTCATCACGTTTGGGATCCTGGGGGCACTGGCTCTCGGCCCTA
TGTGGACCGACCAAGCACTTCACTGAGGCCACCAAGATTGGCCTGTGCCTGTTCTCTCTGG
CCTGCGTGCCCTTTGCCCTGGTGTCCAGCTGCAGGGACAGACCCTTGCCCTGGCTGCCACC
TGCTCGCTGCTCGGGCTGTTTGGCTTCTCGGTGGGCCCGTGGCCATGGAGTTGGCGGTGGA
GTGTTCCCTTCCCGTGGGGGAGGGGCTGCCACAGGCATGATCTTTGTGCTGGGGCAGGCCG
AGGGAATACTCATCATGCTGGCAATGACGGCACTGACTGTGCGACGCTCGAGCCGTCCTTG
TCCACCTGCCAGCAGGGGAGGATCCACTTGACTGGACAGTGTCTCTGTGCTGATGGCCGG
CCTGTGCACCTTCTTCAGCTGCATCCTGGCGGTCTTCTTCCACACCCCATACGGCGCCTGC
AGGCCGAGTCTGGGAGCCCCCTCCACCCGTAACGCCGTGGGCGGCGCAGACTCAGGGCCG
GGTGTGACCGCAGGGGGAGCAGGAAGGGCTGGGGTCTGGGGCCAGCACGGCGACTCCGGA
GTGACCGCGAGGGGGGCTCGCTAGAGGACCCAGAGGGCCCGGAGCCCCACCCAGCCT
GCCACCGAGCGACTCCCGTGCGAAGGCCCAGCAGCCACCGACGCGCCTCCCGCCCCGGC
AGACTCGCAGGCAGGGTCCAAGCGTCCAGGTTTATTGACCGGCTGGGTCTACTCTCTCTT
CTCCTCCCCGTGGGTGATCAGCTAGCTGAGCGCTTGTAGTCCAGGTTGCCCGCCACATCGA
TGGAGGCGAACTGGAACATCTGTTCCACTGCGGGCGGGGCGAAAGGGCTCCTTGGGGCT
CCGGGAGCGAATTACAAGCGCGCACCTGAAAA

FIGURE 169

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50988
><subunit 1 of 1, 560 aa, 1 stop
><MW: 58427, pI: 6.86, NX(S/T): 2
MAGPTEAETGLAEPRALCAQRGHRTYARRWVFLLAISLLNCSNATLWLSFAPVADVIAEDLV
LSMEQINWLSLVYLVVSTPFGVAAIWILSVGLRAATILGAWLNFAGSVLRMVPCMVVGTON
PFAFLMGGQSLCALAQSLVIFSPAKLAALWFPEHQRTANMLATMSNPLGVLVANVLSPLV
KKGEDIPMLGVYTIPAGVVCLLSTICLWESVPTPPSAGAASSTSEKFLDGLKLQLMWNKA
YVILAVCLGGMIGISASFALLEQILCASGHSSGFGSLCGALFITFGIILGALALGPYVDRTK
HFTEATKIGLCLFSLACVPFALVSQIQGQTLALAATCSLLGLFGFSVGPVAMELAVECSFPV
GGAATGMIFVLGQAEGILIMLMTALTVRSEPSLSTCQGGEDPLDWTVSLLLMAGLCTFF
SCILAVFFHTPYRRLQAESGEPPSTRNAVGGADSGPGVDRGGAGRAGVLGPSTATPECTARG
ASLEDPRGPGSPHPACHRATPRAQGPAAATDAPSRPGRLAGRVQASRFIDPAGSHSSFSSPWVIT

Important features:

Signal peptide:

amino acids 1-44

Transmembrane domains:

amino acids 61-79, 98-112, 126-146, 169-182, 201-215, 248-268,
280-300, 318-337, 341-357, 375-387, 420-441

N-glycosylation site.

amino acids 40-43 and 43-46

Glycosaminoglycan attachment site.

amino acids 468-471

FIGURE 171

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48331

<subunit 1 of 1, 1184 aa, 1 stop

<MW: 129022, pI: 5.20, NX(S/T): 5

MMQLLQLLLGLLGGPGGYLFLLGDCQEVTTLTVKYQVSEVPSGTVIGKLSQELGREERRRQA
GAAFQVLQLPQALPIQVDSEEGLLSTGRRLDREQLCRQWDPCLVSFDVLATGDLALIHVEIQ
VLDINDHQPRFPKGEQELEISESASLRTRIPLDRALDPDTGPNLTHTYTLSPSEHFALDVIV
GPDETKHAELIVVKELDREIHSFFDLVLTAYDNGNPPKSGTSLVKVNVLDSDNNSPAFAESS
LALAIQEDAAPGTLLIKLTATDPDQGPNGEVEFFLSKHPPEVLDTFSIDAKTGQVILRRPL
DYEKNPAYEVDVQARDLGNPIPAHCKVLIKVLVDNDNIPSIHVTWASQPSLVSEALPKDSF
IALVMADDLDSGHNGLVHCWLSQELGHFRLKRTNGNTYMLLTNATLDREQWPKYTLTLAQD
QGLQPLSAKKQLSIQISDINDNAPVFEKSRYEVSTRENNLPSLHLITIKAHADLGINGKVS
YRIQDSPVAHLVAIDSNTEGVTQAQRSLNVEEMAGFEFQVIAEDSGQPMPLASSVSVVWSLLDA
NDNAPEVVQPVLSDGKASLSVLVNASTGHLVPIETPNGLGPAGTDTTPPLATHSSRPFLTT
IVARDADSGANGEPLYSIRNGNEAHLFILNPHTGQLFVNVTNASSLIGSEWELEIIVVEDQGS
PPLQTRALLRVMFVTSVDHLRDSARKPGALSMSMLTVICLAVLLGIFGLILALFMSICRTEK
KDNRAYNCREAESTYRQQPKRPQKHQKADIHLVPVLRGQAGEPCEVQGSHKDVDKEAMMEA
GWDPCLOAPFFHLTPTLYRTLNRNQNGGAPAESREVLDQTVNLLFNHPRQRNASRENLNLEP
QPATGQPRSRPLKVAGSPTGRLAGDQGSSEAPQRPPASSATLRQRHLNGKVSPEKESGPRQ
ILRSLVRLSVAFAERNPVEELTVDSPPVQQISQLLSLLHQGQFQPKPNHRGNKYLAKEGGS
RSAIPDTDGPSARAGGQTDPEQEEGPLDPEEDLSVKQLLEBELSSLLDPSTGLALDRLSAPD
PAWMARLSLPLTTNYRDNVISPDAAATEEPRTFTQFGKAEAPELSPTGTRLASTFVSEMSSL
LEMLLEQRSSMPVEAASEALRRLSVCGRTLSDLATSAASGMKVQGDPPGKGTGTEGKSRGSS
SSSRCL

Important features:

Signal peptide:

amino acids 1-13

Transmembrane domain:

amino acids 719-739

N-glycosylation site.

amino acids 415-418, 582-585, 659-662, 662-665 and 857-860

Cadherins extracellular repeated domain signature.

amino acids 123-133, 232-242, 340-350, 448-458 and 553-563

FIGURE 172

CGGACGCGTGGGCGGACGCGTGGGGGAGAGCCGCGAGTCCCGGCTGCAGCACCTGGGAGAAGG
CAGACCGTGTGAGGGGGCCGTGTGGCCCCAGCGTGTGTGGCCCTCGGGGAGTGGGAAGTGGAG
GCAGGAGCCTTCCTTACACTTCGCCATGAGTTTCCTCATCGACTCCAGCATCATGATTACCT
CCCGAGATACTATTTTGGATTGGGTGGCTTTCTTCATGCGCCAAATTGTTTAAAGACTAT
GAGATACGTGAGTATGTTGTACAGGTGATCTTCTCCGTGACGTTTGCATTTTCTTGCACCAT
GTTTGAGCTCATCATCTTTGAAATCTTAGGAGTATTGAATAGCAGCTCCCGTTATTTTCACT
GGAAATGAACCTGTGTGTAATTCTGCTGATCCTGGTTTTCATGGTGCCCTTTTTACATTGGC
TATTTTATTGTGAGCAATATCCGACTACTGCATAAACAACGACTGCTTTTTTCTGTCTCTT
ATGGCTGACCTTTATGTATTTCTTCTGAAACTAGGAGATCCCTTTCCCATTTCTCAGCCAA
AACATGGGATCTTATCCATAGAACAGCTCATCAGCCGGGTTGGTGTGATTGGAGTGACTCTC
ATGGCTCTTCTTCTGGATTGGTGCTGTCAACTGCCATACACTTACATGTCTTACTTCTCT
CAGGAATGTGACTGACACGGATATTCTAGCCCTGGAACGGCGACTGCTGCAAACCATGGATA
TGATCATAAGCAAAAAGAAAAGGATGGCAATGGCACGAGAACATGTTCCAGAAGGGGGAA
GTGCATAACAACCATCAGGTTTCTGGGGAATGATAAAAAGTGTACCATTTCAGCATCAGG
AAGTGAAATCTTACTCTTATTCAACAGGAAGTGGATGCTTTGGAAGAATTAAGCAGGCAGC
TTTTTCTGGAACAGCTGATCTATATGCTACCAAGGAGAGAATAGAATACTCCAAAACCTTC
AAGGGGAAATATTTTAATTTTCTGGTTACTTTTTCTCTATTACTGTGTTTGGAAAAATTTT
CATGGCTACCATCAATATTGTTTTGATCGAGTTGGGAAAACGGATCCTGTCACAAGAGGCA
TTGAGATCACTGTGAATTATCTGGGAATCCAATTGATGTGAAGTTTGGTCCCAACACATT
TCCTTCATTCTTGTGGAATAATCATCGTCACATCCATCAGAGGATTGCTGATCACTCTTAC
CAAGTTCTTTTATGCCATCTCTAGCAGTAAGTCCTCCAATGTCAATTGCTCCTGCTATTAGCAC
AGATAATGGGCATGTACTTTGTCTCTCTGTGCTGCTGATCCGAATGAGTATGCCCTTAGAA
TACCGCACCATAACTACTGAAGTCCTTGGAGAACTGCAGTTCAACTTCTATCACCGTGTGGTT
TGATGTGATCTTCTGGTGCAGGCTCTCTCTAGCATACTCTTCTCTATTGAGCTCACAAAC
AGGCACCAGAGAAGCAAATGGCACCTTGGAACTTAAGCCTACTACAGACTGTAGAGGCCAGT
GGTTTCAAATTTAGATATAAGAGGGGGGAAAAATGGAACAGGGGCTGACATTTTATAAAC
AAACAAAATGCTATGGTAGCATTTTTCACCTTCATAGCATACTCCTTCCCCGTAGGTGATA
CTATGACCATGAGTAGCATCAGCCAGAACATGAGAGGGAGAACTAACTCAAGACAATACTCA
GCAGAGAGCATCCCGTGTGGATATGAGGCTGGTGTAGAGGCGGAGAGGCCAAGAACTAA
AGGTGAAAAATACACTGGAACCTCTGGGGCAAGACATGTCTATGGTAGCTGAGCCAAACAGCT
AGGATTTCCGTTTTAAGGTTACATGGAAAAGGTTATAGCTTTGCCTTGAGATTGACTCATT
AAAATCAGAGACTGTAACAAAAAAAAAAAAAAAAAAAGGGCGGCCGCACTCTAGAGTCG
ACCTGCAGAAGCTTGGCCGCCATGGCCCAACTGTTTTATTGCACTTATAATG

FIGURE 173

MSFLIDSSIMITSQILFFGFGWLFFMRQLFKDYEIRQYVVQVIFSVTFAFSCTMFELIIFEI
LGVLNSSSRYPFHWMNLCVILLVFMVPPYIGYFIVSNIRLLHKQRLLFSCLLWLTFMYFF
WKLGDPPILSPKHGILSIEQLISRVGVIQVTLMALLSGFGAVNCPYTYMSYFLRNVTDTDI
LALERRLLQTMDMIISKKRMMAMARRTMFQKGEVHNKPSGFGWMIKSVTTSASGSENLTLIQ
QEVDALEELSRQLFLETADLYATKERIEYSKTFKGKYNFLGYFFSIYCVWKIFMATINIVF
DRVGKTDPVTRGIEITVNYLGIQFDVKFWSQHISFILVGIIIVTSIRGLLITLTKFFYAIS
SKSSNVIVLLLAQIMGYFVSSVLLIRMSMPLEYRTIITEVLGELQFNFYHRWFDVIFLVSA
LSSILFLYLAHKQAPEKQMAP

Important features:

Signal peptide:

amino acids 1-23

Potential transmembrane domains:

amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398,
425-444

N-glycosylation sites.

amino acids 67-70, 180-183 and 243-246

Eukaryotic cobalamin-binding proteins

amino acids 151-160

FIGURE 174

CATGGGAAGTGGAGCCGGAGCCTTCCTTACACTCGCCATGAGTTTCCTCATCGACTCCAGCA
TCATGATTACCTCCCNGANACTATTTTTTGGATTGGGTGGCTTTTCTTCNGCGCCAATGTT
TAAAGACTATGAGATACGTCAGTATGTTGTACNGGTGATCTTCTCCGTGACGTTTGCCATTT
CTTGCACCATGTTTGGAGCTCATCATCTTTGAAATCTTNGGAGTATTGAATAGCAGCTCCCGT
TATTTTCACTGGAAAATGAACCTGTGTGTAATTCTGCTGATCCTGGTTNTCATGGTGCCTTT
TTACATTGGCTATTTTATTGTGAGCAATATCCGACTACTGCATAAACGACTGCTTTTTT
CCTGTCTCTTATGGCTGACCTTTATGTATTTCCAG

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194

FIGURE 175

GTGTTGCCCTTGGGGAGGGGAAGGGGAGCCNGGCCCTTTCCTAAAAATTGGCCAAGGGTTTC
TTTNTTGAATTCGGGTTNNGNATACCTTCCAGAAAAATATTTTTGGATTGGGGTAGNTT
TTTTTCATGCGCCAATTGTTTAAAGACTATGAGATACGTAGTATGTTGTACAGGTGATNTT
NTCCGTGACGTTTGCATTTTCTGCACCATGTTTGAGCTCATCATNTTTGAAATNTTAGGAG
TATTGAATAGCAGCTCCCGTTATTTTCACTGGAAAAATGAACCTGTGTGTAATTCGTCTGATC
CTGGTTTTTCATGGTGCCTTTTTACATTGGCTATTTTATTGTGAGCAATATCCGACTACTGCA
TAAACAACGACTGCTTTTTTCTGTCTNTTATGGCTGACCTTATGTATTTNTTNTGGAAAN
TAGGAGATCCCTTCCCATCTC

100
90
80
70
60
50
40
30
20
10
0

FIGURE 176

CTCGCGCAGGGATCGTCCCATGGCCGGGGCTCGGAGCCGCGACCTTGGG3GGCTCCGGGATTGTCTACCTTTT
 TGGCTCCCTGCTCGTGAAGTGTCTTTCTCACGGGCTGTGCCTTCAATCTGAGAGTGTAGGGTGCCTTTGCGCAA
 GGAGGGCGAGGCCAGGACCGCTCTTGGCTTCTCTGTGGCCTGCACCGGCAAGTGTAGCCCCGACCCCTGAGCTGTG
 GCTCTGGTGGTGGCTCCCGAGCCCTGGCTCTTCTGGGCAGCAGGCGAATGCATCTGGAGAGCTCTTCAGTGTG
 CCGCTTGTAGCTTGGAGAGAGATGATGCTACTACAGATGGACATCGACAGGGAGAGCTGATATGCAAAAGGAAAGCAA
 GGAGAACCAAGTGGTGGGAGTCAAGTGTTCGGAGCAGGGGCTGGGGGCAAGATTGGTTACCTGTGCACACGATA
 TGAGGCAAGGCAAGGATGGACAGATCCTGGAGACGCGGGGATATGATTTGGTGCCTGTGCTTGTGTCTCAGCCAGGA
 CTATGACCATCCGGGATGAGTTGGATGGTGGGGAAATGGAAGTTCTGTAGGGAGCGCCCGCAAGGCCATGAACAATT
 TGGGTTCTGCAGCAGGGCAGCAGCTGCGCGCTTCTCCCTGTAGCCACTACCTCTCTTTGGGGCCCCGAGAAC
 CTATAATTGGAAAGGCGACGCGCAGGGTGGAGCTCTGTGCACAGGGCTCAGCGGACCTGGCACACCTGGACAGCG
 TCCCTACGAGGCGGGGGGAGAGAAGGAGCAGGACCCCGCTCATCCCGTCCCTGCCAACAGCTACTTTGGCTT
 CTCTATTGACTCGGGGAAAGGTTCTGGTGGTGCAGAAAGAGCTGAGCTTTGTGGCTGGAGCCCCCGCGCCAACCA
 CAAGGGTGTGTGTCTATCTTCGCGCAAGGACAGCGCCAGTCCCTGCTGTGCCAGAGGTTATGCTGTCTGGGGAGCG
 CCTGACCTCCGGCTTTGGCTACTCACTGGCTGTGGCTGACCTCAACAGTGTATGGCTGGCCAGACCTGATATGTGGG
 TGCCCCCTACTTCTTTGAGCGCCAAAGAGAGCTGGGGGGTCTGTGTATGTGTACTTGAACAGGGGGGTCACTG
 GGCTGGGATCTCCCTCTCCGGCTTGGCGCTCCCTGACTCCATGTTTCGGGATCAGCTGGCTGTCTTGGGGGA
 CCTCAACCAAGATGGGTTTCCAGATATTGAGTGGGTGCCCTTTGATGGTGAATGGGAAGCTCTTACATCTACCA
 TGGGAGAGCTGGGGGTTTGTCCCAAACCTTCAAGGTGCTGGAGGCGAGGCTGTGGGATCAAGAGCTTCGGGCTTGG
 CTACTCCCTGTCAAGCAGCTTGGATATGGATGGGAACCAATCCCTGACCTGTGTGGTGGGCTCCCTGGCTGAAC
 CGAGTGGCTCTTCAAGGCGACAGCCATCTCTCATGTCTCCATGAGGTTCTATTGGTCCGAAGCATGCAGCT
 GGAGCAGCCAACTCACTGCTGGCGGCCACTCGGTCTGTGTGGACCTAAAGGTTCTGTTCAGCTACATTTGCACTCCC
 CAGCAGCTATGATCCCTACTGTGGCCCTGGACTATGTTAGATGCGGACAGGCTCCAGGAGGCTCCGGGCGAGGT
 TCCCGTGTGAGCTTCTCGAGCCGTAACCTGGAAAGAACCCAGCACAGGCTCTGGGCAACCGTGTGGCTGAAGCA
 CAGCATGACGAGGATCTGTGGAGCGCATGTTTCCAGCTCCAGGAAATGTCAAAAGCAAGCTTCGGGCACTTGT
 AGTGAATGCTCTGCTACAGTCTCCAGACCCCTGCGCTCCGGCGACAGGCTCTTGGCCAGGGGCTGCTCCAGTGGC
 CCGCATCTCAATGGCCACAGCCAGCCAGCGGCGAGAGATCCATCTCTAGACAGAGGCTGTGTGAAGA
 CAGATCTGCGAGAGCAATCTGCAGCTGTCCAGCGCCGCTTCTGTACCCGGGTCAAGCACACGAAATTTCAACC
 TCTGCCATGGATGTGGATGGAAACAAAGCCCTGTTGCACTGAGTGGGCAAGCACTCATTTGCTGGAGCTGAT
 GGTCAACCACTTGCATCGAGCCAGCCAGCCAGGCTGATGGGATGATGCCATGAAGCCAGCTCTCTGGT
 CATGCTTCTTGACTCATGTCACTACTCAGGGGCTCCGGGCTCTGGACCTGCGGAGAGGCCACTCTGCTGTCCAA
 TGAGAATGCTCCCATGTGTGAGTGTGAGCTGGGGAAACCCATGAAGAGAGTGGCCAGGTCACCTTCACTCAT
 CTTTAGCAGCTCCGGGATCAGCATTTGAGACACGGAACCTGGAGGTAGAGCTGCTGTGTGGCCAGGATCAGTGA
 GAGCTGCATCCAGTCTCTGCACGAGCCGCTGTCTTATTGAGCTGCGACTGTTCATTGTGAGGAATGGGATTTCC
 CCAGCAACTCTTCTTCTGTGTGGTGTGGGGGCGAGAGAGCCATGCAGCTCTGAGCGGGATGTGGGCGAGCAAGGT
 CAAGTATGAGGTCAAGGTTTCCAAACCAAGGCCAGTCCGCTCAGAAACCTGGGCTCTGCTCTTCCCAACATCATGTG
 GCCTCATGAGATTGCCAATGGGAAGTGGTTGCTGTACCAATGCAGGTTGAGCTGGAGGGGCGGCGAGGGGCTGG
 GCAGAAAGGGCTTGTGCTGGAGCCCAACATCTCCACTGGATGTGAAGATGAGGAGTAGAGGCGGCGGGGA
 GCTGGAGCCACCTGAGCAGCAGGAGCTTGTGTAGCGGGCAGGAGCCAGCATGTCTGTGTGGCCAGTGTCTCTG
 TGAGAAAGAGAAAAAATCACTTGGACTGCGCCGGGGCAGCGCAACTGTGTGGTGTTCAGCTGCCCATCTCTA
 CAGCTTTGACCGCGGCTGTGCTGCATGTCTGGGCGCTCTCTGGAACGACACCTTCTGGAGAGGATCTCAG
 TGTGAAGTCCCTGGAGGATGATTGTGGGGCCAAACATCAAGTGAAGTCTCCATAAAGAACTTGATGTCTCGAGA
 TGCCCTCAACAGTGATCCAGTATGGTATATCTTGGACCCCATGGCTGTGGTGGCAGAGGAGTGGCTGTGGGT
 CATCTCTCTGGCTGTACTGCTGGCTGCTGTGGTGTAGCATCTGTGTGCTGTCTGTGGAGATGGGATTTCTT
 CAAACGGGGGAGCACCCGAGGACCACGTCGCCCAAGTACATGCGGTGAAGATTTCTTGGGAAGACGACGACGA
 GTTCAAGAGGAGAGAGCAGGACCAACCTCTGAGGAACATGTGGGCGACCCCGCGGAGAGGCGCCGATGCA
 CCCCATCTGGCTCTGCAGCGGCTCCCGAGCTGGGCGCCGATGGGCACTCAGGCGCAGGACCCGCGCTAGGTTCC
 CATGTCCCACTGCGCTTGTGGTCACTGCCCTCCATCTCTTCCACAGATGGCTCTTGGGATGAAGAGGTATAGT
 GGGCTGTGTGTGTGCTCATCAAGATTGGCAGGATCGGCTCTCTCAGGGGCAACAGCTCTCCACCCCAAGAAC
 TCCCTCCACCAACTCTCCCTTAGAGTGTGTGAGATGAGAGTGGTAAATCAGGACAGGGCCATGGGAGTAGG
 TGAGAAGGGCAGGGGTGTCTGTGCAAAAGGTGGGGAGAAGGGATCCTAATCCCTCTCTCCCATTCACCCGTGT
 GTAACAGGACCCCAAGACATGTGCTCCCGGAAGTGCCTTAACTAGAGGTGGGAGGGGTGTGTGCTACTGA
 CTGAGGCTGTCTCTCTAGTCTTCCCTCTCATCTGACCTTAGTTTGTCTGCTCAGTCTAGTGTGTTCTGTGGT
 TCTCGTATTATTAAAAAATATTTGAGAACAAAAAAGAAAAAAGAAAAA

100791-12403

FIGURE 177

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA55737

><subunit 1 of 1, 1141 aa, 1 stop

><MW: 124671, pI: 5.82, NX(S/T): 5

MAGARSRDPWGASGICYLFGSLLVELLFSRAVAFNLDVMGALRKEGEPGSLFGFSVALHRQL
QPRPQSWLLVGAPQALALPGQQANRTGGFLFACPLSLEETDCYRVDDIDQGADMQKESKENQWL
GVSVRSQGPGGKIVTCAHRYEARQQRVDQILETRDMIGRCFVLSQDLAIRDELDDGGGEWKFCGE
RPQGEHQFGFCQQGTAAAFSPDSHYLLFGAPGTYNWKGATARVELCAQGSADLAHLDDGPYEA
GGEKEQDPRLLIPVPANSYFSGFSDSGKGLVRAEELS FVAGAPRANHKGAVVILRKDSASRLV
PEVMLSGERLTSGFGYSLAVADLNSDGWPDLLVGA PYFFERQEEELGGAVVYVLYNQGGHWAGI
SPLRLCGSPDSMFGISLAVLGDNLQDGFDPDIAGVAPFDGDGKVFIYHGSLSLGVVAKPSQVLE
GEAVGIKSPGYSLSGSLDMDGNQYPDLLVGLSLADTAVLFRARPILHVSHEVSIAPRSIDLEQ
PNCAGGHSVCVDLRVCFYSYIAVPSSYSPTVALDYVLDA DTRRLRGQVPRVTFLSRNLEEPK
HQAGSTVWLKHQHDRVCGDAMFQLQENVKDKLRAIVVTLSYSLQTPRLRRQAPGGQGLPPVAP
ILNAHQPSQTRAEIHFLKQCGGEDKICQSNLQLVHARFCTRVSDTEFQPLPMDVDGTLTALFA
LSGQPVIGLELMVTNLPSDPAQPQADGDDAHEAQLLVMLPDSLHYSGVRALDPAEKPLCLSN
ENASHVECELGNPMKRGAQVTFYILILSTSGISIETTELEVELELLATISBQELHPVVSARARVF
IELPLSIAGMAIPQQLFFSGVVRGERAMQSERDVGSKVKYEVTVSNQGGSLRTLGSFAFLNIM
WPHEITANGKWLLYPMQVELEGGQGGPGQKGLCSRPNNILHLVDSDRDRRRRELEPPEQQEPGE
RQEPSMSWNVPVSSAEKKKNITLDCARGTANCVVFSCLPYSPDRAAVLHVWGRLEWNSSTFLEEY
SAVKSLEIVVRANITVKSSIKNLMLRDASTVIPVMVYLDPMVAVVEGVPWWVILLAVLAGLL
VLALLVLLLWKMGFFKRAKHPEATVPQYHAVKIPREDRQQKFKEEKTGTILRNNWGSPPREGP
DAHPIAADGHPGLGPDGHGPGGTA

Important features:

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 1040-1062

N-glycosylation sites.

amino acids 86-89, 746-749, 949-952, 985-988 and 1005-1008

Integrins alpha chain proteins.

amino acids 1064-1071, 384-408, 1041-1071, 317-346, 443-465, 385-407, 215-224, 634-647, 85-99, 322-346, 470-479, 442-466, 379-408 and 1031-1047

FIGURE 178

CGCGCCGGCGCAGGGAGCTGAGTGGACGGCTCGAGACGGCGCGTGCAGCAGCTCCAGA
AAGCAGCGAGTTGGCAGAGCAGGGCTGCATTTCCAGCAGGAGCTGCGAGCACAGTGCTGGCT
CACAAACAAGATGCTCAAGGTGTGAGCCGTACTGTGTGTGTGTGTCAGCCGCTTGGTGCAGTCA
GTCTCTCGAGCTGCCGCGGGTGGCTGCAGCCGGGGGGCGGTCCGACGGCGGTAATTTTC
TGGATGATAAAACAATGGCTCACCAAACTCTCAGTATGACAAGGAAGTCGGACAGTGGAAAC
AAATTCGAGACGAAGTAGAGGATGATTATTTCCGCACCTGGAGTCCAGGAAAACCCCTTCGA
TCAGGCTTTAGATCCAGCTAAGGATCCATGCTTAAAGATGAAATGTAGTCGCCATAAAGTAT
GCATTGCTCAAGATTCTCAGACTGCAGTCTGCATTAGTCACCGGAGGCTTACACACAGGATG
AAAGAAGCAGGAGTAGACCATAGGCAGTGGAGGGGTCCCATATTATCCACCTGCAAGCAGTG
CCCGATGGTCTATCCAGCCCTGTTTGTGGTTCAGATGGTCATACCTACTCTTTTTCAGTGCA
AACTAGAATATCAGGCATGTGTCTTAGGAAAACAGATCTCAGTCAAATGTGAAGGACATTGC
CCATGTCCCTTCAGATAAGCCCACCAAGTACAAGCAGAAATGTTAAGAGAGCATGCAGTGACCT
GGAGTTCAGGGAAGTGGCAAAACAGATTGCGGGACTGGTTCAGAGGCCCTTCATGAAAGTGGAA
GTCAAAAACAAGAAGCAAAAACATTGCTGAGGCCTGAGAGAAGCAGATTGATACACAGCATC
TTGCCAATTTGCAAGGACTCACTTGGCTGGATGTTTAAACAGACTTGATACAACTATGACCT
GCTATTGGACCAAGTCAGAGCTCAGAAGCATTTTACCTTGATAAGAAATGAACAGTGTACCAAGG
CATTCTTCAATCTTGTGACACATACAAGGACAGTTTAAATATCTAATAATGAGTGGTGCTAC
TGCTTCCAGAGACAGCAAGACCCACCTTGCCAGACTGAGCTCAGCAATATTCAGAAGCGGCA
AGGGGTAAAGAAGCTCCTAGGACAGTATATCCCCCTGTGTGATGAAGATGGTTACTACAAGC
CAACAACAATGTGATGGCAGTGTGGACAGTGTGGTGTGTTGACAGATATGAAATGAAGTC
ATGGGATCCAGAATAAATGGTGTTCAGATTGTGCTATAGATTTTGAGATCTCCGGAGATTT
TGCTAGTGGCGATTTTTCATGAATGGACTGATGATGAGGATGATGAAGACGATATTATGAATG
ATGAAGATGAAATGAAGATGATGATGAAGATGAAGGGGATGATGATGATGGTGGTATGAC
CATGATGTATACATTGATGATGACAGTTGAAATCAATAAATCTACATTTCTAATATTTA
CAAAAATGATAGCCTATTTAAAATTATCTTCTCCCCAATAACAAAATGATTCTAAACCTCA
CATATATTTGTATAATTATTTGAAAAATTGCAGCTAAAGTTATAGAAGTTTATGTTTAAAT
AAGAATCATTTGCTTTGAGTTTTTATATTCTTACACAAAAGAAAATACATATGCAGTCTA
GTCAGACAAAAATAAGTTTTGAAGTGCTACTATAATAAATTTTTCACGAGAACAACTTTGT
AAATCTTCCATAAGCAAAATGACAGCTAGTGCTTGGGATCGTACATGTTAATTTTTTGAAAG
ATAATCTAAGTGAAATTTAAAATAAAATAAATTTTTAATGACCTGGGTCTTAAGGATTTAGG
AAAAATATGCATGCTTTAATTGCATTTCCAAAGTAGCATCTTGCTAGACCTAGATGAGTCAG
GATAACAGAGAGATACCACATGACTCCAAAAAAAAAAAAAAAA

FIGURE 179

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49829
><subunit 1 of 1, 436 aa, 1 stop
><MW: 49429, pI: 4.80, NX(S/T): 0
MLKVSAVLCVCAAAWCSQSLAAAAVAAAGGRSDGGNFLDDKQWLTTISQYDKEVGQWNKFR
DEVEDDYFRTWSPGKPFQALDPAKDPCLMKKCSRHKVCIAQDSQTAVCISHRRLTHRMKEA
GVDHRQWRGPILSTCKQCPVVPSPVCGSDGHTYSFQCKLEYQACVLGKQISVKCEGHCPCP
SDKPTSTSRNVKACSDLEFREVANRLRDWFKALHESGSQNKKTKTLRLPERSRFDTSILPI
CKDSLGMFMNRLDTNYDLLLDQSELRSIYLDKNEQCTKAFNNSCDTYKDLSISNNEWCYCFQ
RQQDPPCQTELSNIQKRQGVKKLLGQYIPLCDEDGYKPTQCHGSVGCWCVDRYGNEVMGS
RINGVADCAIDFEISGDFASGDFHEWTDDEDEDDIMNDEDEIEDDEDEGDDDDGGDDHDVYI
```

Important features:

Signal peptide:

amino acids 1-16

Leucine zipper pattern.

amino acids 246-267

N-myristoylation sites.

amino acids 357-362, 371-376 and 376-381

Thyroglobulin type-1 repeat proteins

amino acids 353-365 and 339-352

FIGURE 180

CAGACTCCAGATTTCCTGTCAACCCAGGAGTCCAGAGAGGAAACGCGGAGCGGAGACAACAGTACCTGAGCC
CTCTTTTCAGCCCGGGATCGCCCCACAGGGCA~~ATGGCGCA~~ACAGATCTGGCTGCCCTCCCCGGTGCTCTCTTGCC
GCTCTGCCCTCCGGTGCTGCTGCCCTGGGCGGGCGGCTTCACACCTTCCTCGATAGCGACTTCACCTTTACCTCT
CCCCCGCGGCAAGGAGTGTCTTCTACACGCCCATGCCCCGGAAGGCTCGCTGGAGATCGAGTACCAAGTTTGA
GATGGAGCAGGATAGATATGTGATTTCCATCTTGCTCTCTCAGAAGGCAAAACCTTAGTTTGTGAACAAGAAAA
TCAGATGAGTTCACACTGTGAGACTGAAAGTTGGTGATTACATGTTCTGCTTTGACAACTTCAGACCCATT
TCTGAGAAGGGTATTTCTTGGAATTAATCCTGGATAATATGGGAGAACAGGCACAGAAACAAGAAATTTGGAAG
AAATATATTACTGGCACAGATATTTGGATATGAAACTGGGAAGACATCTCGGAATCCATCAACAGCATCAAGTCC
AGACTAAGCAAAAGTGGGCACATACAAATCTGCTTAGAGCATTGGAAGCTCGTGATCGAAACATACAGAAAGC
AACTTTGATAGAGTCAATTTCTGGTCTATGGTTAATTTAGTGGTCATGGTGGTGGTGCAGCCATTCAAGTTTAT
ATGCTGAAGAGTCTGTTTGAAGATAAGAGGAAAGTAGAACT~~TAA~~AACCTCCAACTAGAGTACGTACATTTGAAA
AATGAGGCATAAAAAATGCAATAAATCTGTACAGTCAAGACCATTAATGGTCTTCTCCAAATATTTTGGATATA
AAAGTAGGAAACAGGTATAATTTTAATGTGAAAAATTAAGTCTTCACTTTCTGTGCAAGTAATCTGCTGATCCAG
TTGTACTTAAAGTGTGTAACAGGAATATTTTGAGAATATAGGTTTAACTGAATGAAGCCATATTAATAACTGCAT
TTTCCCTAACTTTGAAAAATTTTGCAAAATGTCTTAGGTGATTTAAATAAATGAGTATTTGGGCCTAATTTGCAACACC
AGTCTGTTTTAAACAGGTTCTATTACCCAGAACTTTTGTAAATGCGGCAGTTTACAAATTAACCTGTGGAAGTTT
TCAGTTTAAAGTTATAAATCACTGAGAAATTACCTAATGATGGATGTAATAATCTTTAGAGTACAAAGCCCAA
CTTTTCTCTATTATACATATGCTCTCTCTATAATGTAAATAGAATAATAGCTTTGAAATCAATATAGTTTTCG
AGATTTTATAACCAAATACATTTTCAGTGTAAACATATAGCAAGACATTTAGTCTTTGTACTTTGCTTACATTT
CCAAAAGCTGACATTTTCACGATTTCTTAAAAACAAGAGTTACACTTACTAAATTAGGACATGTTTCTCTTTG
AAATGAAGAAATATAGTTTAAAGAGTCTCTCTCTCAATAGGCAACATTTTCTCTAACCTTAACTAAAGGTAGGA
TTTTAAATTTAAATGTGAGGTAAATTAAGTTTATTTTAAATAGTATCTGTCAGTTAAATTAATCTGTCACAGTTAA
TAATCATGTTATGTAAATTTTAAATCATGATGTGCTGACTTGGATAATTCATTTTACCAGCAGTTATGAAGGAAATA
TTGCTAAATGATCTGGGCCCTACCATAAATAAATATCTCTTTCTGAGCTCTAAGAAATTACAGAAAACAGGAA
AGAATTTAGAAAACCTTGAGAAAACCTTAATCCAAAATAAAATTCATTTAAAGTATGAAATATAAATATCTAGA
ATCTGACTGGCTCATCTGACATCTCTACTCATAACATAAATCAAAGGAGATGATTAATTTCCAGTTAGCTGGAAG
AACTTTGGCTGTAGGTTTATTTATTTTCAAGAAATTCGGTTTGAATTTTGTGAAGCAGGTACATTTTATATA
AAATGTAGGCCCTACTGTAAAGTTTACGACTGGGTGTACATATTTATAAAAATTTTATTATAACCAACTTTTAT
TAAAAATGGCCCTTCTGAAACCTTTATTTATGATGTTGAAGTAAGGATTAGAAAACATAGACTCCCAAGTTTAAAA
CACTTAAATGTGAATAACCCATATACAAACAAAGTTTCTGCCATCTAGCTTTTGAAGTCTATGGGGGTCTTAC
TCAAGTACTAGTAAATTAACCTATCATGATGAATGAACTATAAATTTTAAAGTATGCCATTTATAAGTTATGTTTAT
GACTACATTTGAGTTAGAAACCAAGCTTAAAAATTTGGGGTATAGAACCCCTCAACAGGTTAGTAATGCTGGAATTT
CTTGATGAGCAATAATGATAACAGAGAGTGATTTCAATTTACACTCATAGTAGTATAAAAAGAGATACATTTCCC
CTTTAGGCCCTCGGAGAGAGCAGCTTAGATTTCCCTACTGGCAAGGTCTTTAAAAATGAGGTAAATGCCGAT
ATGATCAATTAACCTTAATTTGGCCCAAGAAATGCTTCAGGTGCTAGGGGTATTTTCAACACTTGCAGACCTTGCAGAACAA
AGGTCAATAAGATCTTGCTTATGAATACCCCTCCCTTTTGGCGTGTAAATTTGCAATGAGAAGCAAAATTACAA
ACATGTATGATTTTGTGCCATGATCTTAAACCTATGATTCAGTAACCTTTACCATATAAAAAAGATAATTTGCTT
TATTTTGAAGAAGAAATTTAGGAATCAATAGGACAATTAATTTTATAGCAAGGTAAAAAGACAGATATTTAAGAGG
CATAAACAAAAGCAAACTTGTAAACAGAGTAAAAATCTTTAATATTTTAAAGACATATCTGTTATCTGCTT
CATATGCTTTTAAATTTCACTTATCCATTTTAAATTAAGTTATGCTTAAATTTAGTATGAGTAACTGTTTATCACTT
AACAGCTCAATTTTGTCTTTTCAATATACAAATTTTAAAAATACTACAAATTTTAACTAAGGCCCAACCGATTTT
CATATAGTAGCAGTTACCGGTTCACCTCACATAAGGCCTAGAGTTTGTCTGATATGCATTTGGATGATTAAT
GTATGCTGTTCTTCAATGATGTTCAAGACATGGAGGGTGTGTGAATTTTATGGTAAAAATTTAAGTAAATCTCTTCA
CACATAATGCTGCTTAAAAATGCAACAAAATGAGCACTTCAATTTGATGTCTCTCTCAATGAGATTTCTTAT
GTGAAATTTTAAAGACATTTGATTCGCATGTAAAGATTTTTCATCTGAAGTATCAATTAAGTCAATCAGTGTGT
CTCAACTGCTTTATACTTATAACAGCACTCTTAAATAAGCAACGTATTTGAGTACTGATATGATATAATAA
AAATTTCAAGAGGAAAA

003337.12401

FIGURE 181

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52196
><subunit 1 of 1, 229 aa, 1 stop
><MW: 26017, pI: 4.73, NX(S/T): 0
MGDKIWLPFFVLLLAALPPVLLPGAAGFTPSLDSDFTLTPAGQKECFYQPMPLKASLEIEY
QVLGDAGLDIDFHLASPEGKTLVFEQRKSDGVHTVETEVGYMFCFDNTFSTISEKVIFFEL
ILDNMGEQAQEEDWKKYITGTDILDMKLEDILESINSIKSRLSKSGHIQILLRAFEARDRN
IQESNFRVNFWSMVNLVVMVVVSAIQVYMLKSLFEDKRKRSRT
```

Important features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 195-217

N-myristoylation site.

amino acids 43-48

Tyrosine kinase phosphorylation site.

amino acids 55-62

FIGURE 182

CCATCCCTGAGATCTTTTTATAAAAAACCCAGTCTTTGCTGACCAGACAAAGCATACCAGAT
CTCACCAGAGAGTCGCAGACACTATGCTGCCCTCCCATGGCCCTGCCAGTGTGCTGCGATG
CTGCTTTCCCTGCCCTCATTCTCCTGTGTCAAGGTGAAGAAACCCAGAAGGAACTGCC
CTCTCCACGGATCAGCTGTCCCAAAGGCTCCAAGGCCATGGCTCCCCCTGCTATGCCTTGT
TTTTGTACCAAAATCCTGGATGGATGCAGATCTGGCTTGCCAGAAGCGGCCCTCTGGAAAA
CTGGTGTCTGTGCTCAGTGGGGCTGAGGGATCCTTCGTGTCTCCCTGGTGAGGAGCATTAG
TAACAGCTACTCATACATCTGGATTGGGCTCCATGACCCACACAGGGCTCTGAGCCTGATG
GAGATGGATGGGAGTGGAGTAGCACTGATGTGATGAATTACTTTGCATGGGAGAAAAATCCC
TCCACCATCTTAAACCCCTGGCCACTGTGGGAGCCTGTCAAGAAGCACAGGATTTCTGAAGTG
GAAAGATTATAACTGTGATGCAAAGTTACCCATGTCTGCAAGTTCAAGGACTAGGGCAGGT
GGGAAGTCAGCAGCCTCAGCTTGGCGTGCAGCTCATCATGGACATGAGACCAGTGTGAAGAC
TCACCCCTGAAGAGAATATTCTCCCAAACCTGCCCTACCTGACTACCTTGTGATGATCCTCC
TTCTTTTCTTTTCTTTCACCTTCATTTCAGGCTTTTCTCTGTCTTCCATGTCTTGAGATC
TCAGAGAATAATAATAAAATGTTACTTTATAAAAAAAAAAAAAAAAAAAAAA

FIGURE 183

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56965

<subunit 1 of 1, 175 aa, 1 stop

<MW: 19330, pI: 7.25, NX(S/T): 1

MLFPMALPSVSWMLLSCLILLCQVQGEETQKELPSPRISCPKGSKAYGSPCYALFLSPKSWM
DADLACQKRPSGKLVSVLGAEGSFVSSLVRSISNSYSYIWIGLHDPTQGSEPDGDGWEWSS
TDVMNYFAWEKNPSTILNPGHCGSLSRSTGFLKWKDYNCDAKLPYVCKFKD

Important features:

Signal peptide:

amino acids 1-26

C-type lectin domain signature.

amino acids 146-171

FIGURE 184

CCAGTCTGTCGCCACCTCACTTGGTGTCTGCTGTCCCCGCCAGGCAAGCCTGGGGTGAGAGC
ACAGAGGAGTGGGCCGGGACCATGCGGGGGACGCGGCTGGCGCTCCTGGCGCTGGTGTGGC
TGCCCTGCGGAGAGCTGGCGCCGGCCCTGCGCTGCTACGTCTGTCCGGAGCCACAGGAGTGT
CGGACTGTGTCACCATCGCCACCTGCACCACCAACGAAACCATGTGCAAGACCACACTCTAC
TCCCGGGAGATAGTGTACCCCTTCCAGGGGGACTCCACGGTGACCAAGTCCTGTGCCAGCAA
GTGTAAGCCCTCGGATGTGGATGGCATCGGCCAGACCCTGCCCGTGTCTGCTGCAATACTG
AGCTGTGCAATGTAGACGGGGCGCCCGCTCTGAACAGCCTCCACTGCGGGGCCCTCACGCTC
CTCCCACTCTTGAGCCTCCGACTGTTAGAGTCCCCGCCACCCCCATGGCCCTATGCGGCCCA
GCCCCGAATGCCTTGAAGAAGTGCCCCCTGCACCAGGAAAAAAAAAAAAAAAAA

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181

FIGURE 185

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56405

<subunit 1 of 1, 125 aa, 1 stop

<MW: 13115, pI: 5.90, NX(S/T): 1

MRGTRLALLALVLAACGELAPALRCYVCPEPTGVSDCVTIATCTTNETMCKTTLYSREIVYP
FQGDSTVTKSCASKCKPSDVGIGQTLFVSCNTELCNVDGAPALNSLHCGALTLLPLLSLRL

Important features:

Signal peptide:

amino acids 1-17

N-glycosylation site.

amino acids 46-49

FIGURE 186

CTGCAGTCAGGACTCTGGGACCGCAGGGGGCTCCCGGACCCTGACTCTGCAGCCGAACCGGC
ACGGTTTCGTGGGGACCCAGGCTTGCAAAGTGACGGTCATTTTCTCTTTCTTCTCCCTCTT
GAGTCCTTCTGAGATGATGGCTCTGGGCGCAGCGGAGCTACCCGGGTCTTTGTGCGGATGG
TAGCGGCGGCTCTCGGGGGCCACCCTCTGCTGGGAGTGAGCGCCACCTTGAATCGGTTCTC
AATTCCAACGCTATCAAGAACCTGCCCCACCGCTGGGCGGCGCTGCGGGGACCCAGGCTC
TGCAGTCAGCGCCGCGCGGGAATCCTGTACCCGGGCGGGAATAAGTACCAGACCAATTGACA
ACTACCAGCCGTACCCGTGCGCAGAGGACGAGGAGTGCGGCACTGATGAGTACTGCGCTAGT
CCCACCGCGGAGGGGACGAGGCGTGCAAACTGTCTCGCTGCAGGAAGCGCCGAAAAAG
CTGCATGCGTCACGCTATGTGCTGCCCGGGAATTACTGCAAAAATGGAATATGTGTGCTT
CTGATCAAAATCATTTCCGAGGAGAAATTGAGGAAACCATCACTGAAAGCTTTTGGTAATGAT
CATAGCACCTTGGATGGGTATTCCAGAAGAACCACCTTGTCTTCAAAAATGTATCACACCAA
AGGACAAGAAGGTCTGTITTTGTCTCCGGTCATCAGACTGTGCCTCAGGATTGTGTTGTGCTA
GACACTTCTGGTCCAAGATCTGTAAACCTGTCTGAAAGAAGGTCAAGTGTGTACCAAGCAT
AGGAGAAAAGGCTCTCATGGACTAGAAATATTCCAGCGTTGTTACTGTGGAGAAGGTCTGTCT
TTGCCGGATACAGAAAGATCACCATCAAGCCAGTAATTTCTCTAGGCTTCACACTTGTGAGA
GACACTAAACCAGCTATCCAAATGCAGTGAATCCTTTTATATAATAGATGCTATGAAAACC
TTTTATGACCTTCATCAACTCAATCCTAAGGATATACAAGTTCTGTGGTTTCAGTTAAGCAT
TCCAATAACACCTTCCAAAACCTGGAGTGTAAGAGCTTTGTTTCTTTATGGAACCTCCCTGT
TGATTGCAGTAAATTACTGTATTGTAAATCTCAGTGTGGCACTTACCTGTAAATGCAATGA
AACTTTTAATTATTTTCTAAAGGTGCTGCACTGCCATTTTTCTCTTGTATGTAAATTT
TTGTACACATTGATTGTTATCTTGACTGACAAATATTCTATATTGAAGTAAATCAAT
TCAGCTTATAGTTCTTAAAGCATAACCTTTTACCCATTAAATTCTAGAGTCTAGAACGCA
AGGATCTCTTGAATGACAAATGATAGGTACCTAAATGTAACATGAAAATACTAGCTTATT
TTCTGAAATGTACTATCTTAATGCTTAAATTATATTCCCTTTAGGCTGTGATAGTTTTTGA
AATAAAATTTAACATTTAAAAA

FIGURE 187

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57530

<subunit 1 of 1, 266 aa, 1 stop

<MW: 28672, pI: 8.85, NX(S/T): 1

MMALGAAGATRVFVAMVAAALGGHPLLGVSATLNSVLNSNAIKNLPPLGGAAGHPGSAVSA
APGILYPGGNKYQOTIDNYQYPYCAEDEECGTDEYCASPTRGGDAGVQICLACRKRKRKCMRH
AMCCPGNYCKNGICVSSDQNHFRGEIEBETITESFGNDHSTLDGYSRRTTLSSKMYHTKGQEG
SVCLRSSDCASGLCCARHFWSKI CKPVLKEGQVCTKHRRKGSHGLEIFQRCYCGEGLSCRIQ
KDHHQASNSSRLHTCQRH

Important features:

Signal peptide:

amino acids 1-23

N-glycosylation site.

amino acids 256-259

Fungal Zn(2)-Cys(6) binuclear cluster domain

amino acids 110-126

FIGURE 188

TGTGTTTCCCTGCAGTCAGAATTTGGGACNGCAGGGGTTCCCGGACCTGATTTTGCAGCGGA
ACGGGAAGGTTTTGTGGGACCCAGGTTGAAATGACGGTCATTTTTTTTCTTTCTCCTTCNG
GAGTCCTTNTGAGANGATGGTTTTTGGGCGCAGCGGGAGCTAACCCGGTTTTTTGTNGCGATG
GTAGCGGCGGTTTTTCGGCGGCCACCTTNTGCTGGGAGTGAGCGCCACCTTGAATCGGTTTTTC
AATTCCAACGNTATCAAGAACCTGCCCCACCGNTGGGCGGCGCTGCGGGGCACCCAGGNTT
TGCAGTCAGCGCCGCGCCGGGAATCCTGTACCCGGGCGGGAATAAGTACCAGACCATTGACA
ATTACCAGCCGTACCCGTGCGCAGAGGACGAGGAGTGCGGCACTGATGAGTACTGCGCTAGT
CCCACCCGCGAGGGGGANGCGGGCGTGCAAATNTGTTNTNGCCTGCAGGAAGCGCCGAAAACG
CTGCATGCGTCANGCTATGTGCTGCCCCGGGAATTACTGCAAAAATGGAATATGTGTGNTT
CTGATCAAAATCATTTCCGAGGAGAAATTGAGGAAACCATCACTGAAAGCTTTGGTAATGAT
CATAGCACCTTGGATGGG



FIGURE 189

GAGGAACCTACCGGTACCGGCCGCGCGCTGGTAGTCGCCGGTGTGGCTGCACCTCAACCAATCCCGTGGCCGGG
CTGGGGCCCTGGAGAGTGGCTGTGTCTTCTCTCTGCACGCCGTGGCTTGGGCTCGGCCAGGGCGGGTCCGCCGCCA
GGGTCTGGAGATGGGGAGTAGCTACAGGAAGCGACCCGCGATGGCAAGGTATATTTTGTGGAATGAAAAGG
AGTATTAGAAAATGAGCTGAAGACCATTACAGAGTTAAATATTTTGGGGACAGATTTTGTGATGCTTGATTCACCTT
TGAAGTAATGTAGACAGAAGTTCTCAAATTTGCATATTACATCACTGGAAAGCAGCAGTGAAATCTTAATGTTTAC
TTAAATCAGAACTTGCAATAGAAAGAGAATGGGAGTCTGGTTAAATAAAGATGACATATACAGAGACTTGAAAAG
GATCATTTCTCTGTTTTCTGATAGTGTATATGGCCATTTTGTGGGACAGATCAGGATTTTACAGTTTACTTGG
AGTGTCTCAAAACTGGCAAGCTAGAGAAATAAGACAAGCTTTCAAGAAATTTGGCATTGGAAGTTACATCTCGATAA
AAACCAGGAATACCCAAATGACATGCGGATTTTTTAAATAATAATAGAGCATATGAAGTACTCAAGATGAAGA
TCTACGGAAAAAGTATGACAAATATGGAGAAAAGGAGCTTGAGGATAATCAAGGTGGCCAGTATGAAAGCTGGAA
CTATTATCCTTATGATTTTGGTATTTATGATGATGATCCTGAAATCATAACTTTGAAAGAGAGAAATTTGATGC
TGCTGTTAATCTGGAGAACTGTGGTTTGTAAATTTTACTCCCCAGGCTGTTACACACTGCCATGATTTTAGCTCC
CACATGGAGAGACTTTGCTAAAGAAAGTGGATGGGTACTTTCGAATTTGGAGCTGTTAACTGTGGTGTATGATGAAT
GCITTTGCCGAATGAAAGGAGTCAACAGCTATCCCAGTCTCTTCATTTTTCGGTCTGGAATGGCCCCAGTGAATA
TCATGGAGACAGATCAAGAGGAGTTTTAGTGAGTTTTCGAATGCAGCATGTTAGAAATCAGTGACAGAACTTTG
GACAGGAAATTTTGTCAACTCCATCAAACTGCITTTTGTCTGGTATTTGGCTGGCTGATCACTTTTGTCTCAA
AGGAGGAGATTTTGTGACTCTCAGACACAGCATCAGGCTTAGTGGCATGTTGTTTCACTCATTTGGATGCTAA
AGAAATATATTTGGAAGTAAATCAATAATCTCCAGATTTTGAACCTACTTTCGGCAAAACACATCAGAGGATCGTT
GGCTCATCACTCGGTGGCTGTATTTTTCATTTTGGAAAAATAGAAATCTCAAATGATCCTGAGCTGAAAAAATG
AAAAACTCTACTTAAAAATGATCATATTCAGTTGGCAGGTTTGACTGTTCTCTGCACAGCATCTCGATGATA
TCGTATGTTTTCAGCCGCTCTAGCAGTATTTAAAGGACAAGGAACCAAGAATATGAAATTTGATCTGGAA
GAAGATTTCTATGATATCTTGCTTTGCGAAAGAAAGTGAATCTCATGCTTACCACTGTCAGCTCAAAA
TTTTCTGCGAATGACAAAGAAACATGGCTTGTGATTTCTTTGCCCTGGTGGTCCACCATGTCAGGCTTTACT
ACAGAGTTTCCAGAGACATCAAACTCTTTTATCCGTCAGCTTAAGTTTGGTACATAGATTGTACAGTTCTAG
GGCATCTGTAACTATGATTAACAACTCAGGCTTATCCAACACAGTGGTATTAACACAGTCCAATCTCATGAT
TGAAGGACATCACTCTGTCGAACAAATCTTGGAGTTACATAGAGATCTTATGAATCCTTCAGTGGTCTCCCTTAC
ACCCACCACTTCAACGAATCTAGTTTACAAAGAAACACAACGAAGTCTGGATGTTGATTTCTATTCTCCGCTG
GTGTCATCCTTGCCAAGTCTTAAGTCCGAAATGGAAAGAAATGAGCATCAATGATCTTGAATCAGTCAACGCTGGG
CAGTATAGATTGCCAACAGTATCATCTTTTGTGCCAGGAAAACGTTCAAAGATCACTTGAGATTAAGATTTT
TCCCCAAATCAAAATAAGCTTTATCAGTATCAGTATCAATAGTGGTGGAAATGGGATGCTTATCCCTGAGAAT
CTGGGCTCAGGATTTTATCTCAAGTATCCAAGATCTAACCTCAGACTTTCAGTGAAAAAGTTTCTACAGG
GAAAAATCACTGGGTGATGATTTCTATGCTCCTTGGTGTGGACCTGCCAGAAATTTGCTCCAGAAATTTGAGCT
CTTGCTAGGATGATTAAGGAAAGTGAAGCTGAAAGCTGAAAGTAGACTGTGAGGCTTATGCTCAGACATGCCAGAA
AGCTGGGATCAGGCTCTTCAACTGTTAAGTTTATTTCTAGAAAGAGCAAGAGAAATTTTCAAGAAAGAGCA
GATAAATACCAAGATGCAAAAGCAATCGCTGCCTTAATAAGTGAATAATTTGAAACTCTCCGAAATCAAGGCAA
GAGGAATAGAGTGAACCTTGAATATGTTGAAGTAGAAGAAAAAGTTTAAAGAAATTTCTGACAGATGACATCAG
AAGACACCTATTAGAAATGTTACATTTATGATGGGAATGAATGAACATTTCTTAGACTTGCAGTGTGATCTGCCA
GAATTTATCTACAGCATCTGGTGTAAAGAGAGGCTCTGCAAACTTTTCTGTAAAGGGCCGGTTATAAATATTTTGA
GACTTTTGAGGCTATAATATATGTTTCAACATGAGAAACAGAAATAGAGTCATCATGTTATCTTTTGTATTTGCT
TTTAAACAACTTTAAAAAATTTAAAAAGATTCTTAGCTCAGAGCATAACAAAGTAGGCTGGATTCAGTCCATG
GAOCTAGAGTTGCTGTCCCCCTCGAAGGACTTATAATGTTTTCAGGTGGCTGAGCTGAACATGAGTCTGTGCTGT
ATCTCATTAATATGTTCAAGTTGTATAAAGTCCACTTCCCTTCAAGTTTTCAGTGTGACTGAAAGAGGTAACCT
TAGTTTTTGTGCTGTTTCTTCAAAAATGCTATCCCTAACCATATATTTATTTTCGTTTAAAAACACCACT
GATGTGGCAACAGTAAACAAACCTGTTATGCTGTATTTATGAGGAGATTTCTCATTTGTTTTCTTCTCTTCA
AAGGTTGAAAAAATGCTTTTAAATTTTCAAGCCGAGAAACAGTGCAGCAGTATGTGCACAGATGAATACCA
AAATTTGAGCAACAGTAAGTGCACAAATCTGTAGTTTGTGTATCATCCAGGAAAACTGAGGGAAGAAAAATTA
TAGCAATTAACCTGGGATGTAGAGATCTCTAAATATGTTATCAAGTATTTAGAGTTCTATATTTTAAAGATATA
TGTGTTCAATGATTTTCTGAAATGTCTTCATAGAAATTTCCCACTGATAGTTGATTTTGGAGGATCTAATAT
TACATATTTTGGCTCTGCAAGTTTGTGACCTGTATCCTTTATTTACATTTGGGTTTTCTTTCATAGTTTGG
TTTTCTCCTCTGTGTCAGTCTATTATTTCAATAGGAAAAATTTACTTTACAGGTTGTTTTCTGCTAGCTTAT
AATGATCATGTAGTTTATCCAGTTTACTAGTTTACTGTGACAGGGCTGCCITTTTCAGATAAATTTGACATAATA
ACTGAAGTTATTTTATAAGAAATCAAGTATATAAGTCTAGGAAAGGGATCTCTGAGTTTCTGTGTTGTTTGA
CTCAAGAAATCAAAATTTGTCAGTAACTGTAGTTTGTAGTTATTAATCAGAGTGCAGAGATGGTAAAAAT
CCAATCAGTCAAAGAGGTCATGAATTTAAAGGCTTGCAACTTTTCAAAAAAATAAAAAA

FIGURE 190

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56439

<subunit 1 of 1, 747 aa, 1 stop

<MW: 86127, pI: 7.46, NX(S/T): 2

MGVVLNKKDDYIRD LKRIILCF LIVYMAILVGT DQDFYSL LGVSKTASSREIRQAFKKLALKL
HPDKNPNNPNNAHGDFL KINRAYEVLKDEDLRKKYDKYGEKGLDNQGGQYESWNYRYDFGI
YDDDEPEIITLERREFDA AVNSGELW FVNFYSPGCSHCHDLAPTWRDFAKEVDGLLRIGAVNC
GDDRMLCRMKGVNSYPSLFIFRSGMAPVKYHGDRSKESLVSFAMQHV RSTVT ELWTGNFVNS
IQTAFAAGIGWLITFC SKGGDCLTSQTRLRLSGMLFLNSLDAKEIYLEVIHNLPDFELLSAN
TLEDRLAHHRWLLFFHFGKNENSNDPELKKLKTLLKNDHIQVGRFDCSSAPDICS NLYVFQP
SLAVFKGQGTKEYEIH HGKILYDILAFAKESVNSHVTTLG PQNF PANDKEPWLVDFFAPWC
PPCRALLPELRRASNLLYQ LKFGTLDCTVHEGLCNMYNIQAYPTTVVFNQSN IHEYEGHHS
AEQILEFIEDLMNP SVVSLTPTTFNELVTQRKHNEVWMVDFYSPWCHPCQVLMPEWKRMART
LTGLINVGSIDCQQYHSFCAQENVQRYPEIRFFPPKSNKAYQYHSYNGWNRDAYS LRIWGLG
FLPQVSTDLTPQTFSEKVLQGNHWVIDFYAPWCGPCQNFAP EPELLARMIKGKV KAGKVDC
QAYAQCQKAGIRAYPTVKFYFYERAKRNFQEEQINTRDAKATAALISEKLET LNRNQGRKNKDEL

Important features:

Endoplasmic reticulum targeting sequence.

amino acids 744-747

Cytochrome c family heme-binding site signature.

amino acids 158-163

Nt-dnaJ domain signature.

amino acids 77-96

N-glycosylation site.

amino acids 484-487

[illegible]

AGACAGTACCTCCTCCCTAGGACTACACAAGGACTGAACCAGAGAAGGAAGGACAGAGCAATTA
GCCATGAACATCATCCTAGAAATCCTTCTGCTTCTGATCACCATCATCTACTCCTACTTGGAA
GTCGTTGGTGAAGTTTTTTCATTCCTCAGAGGAGAAAATCTGTGGCTGGGAGATTGTTCTCA
TTACTGGAGCTGGGCATGGAATAGGCAGGCAGACTACTTATGAATTGTCAAAACGCAGAGAC
ATATTGGTTCTGTGGGATATTAATAAGCGCGGTGTGGAGGAAATGCAGCTGAGTGCCGAA
ACTAGGCGTCACTGCGCATGCGTATGTGGTAGACTGCAGCAACAGAGAAGAGATCTATCGCT
CTCTAAATCAGGTGAAGAAAGAAAGTGGGTGATGTAACAATCGTGGTGAATAATGCTGGGACA
GTATATCCAGCCGATCTTCTCAGCACCAAGGATGAAGAGATTACCAAGACATTTGAGGTCAA
CATCCTAGGACATTTTTGGATCAGAAAAGCACTTCTTCCATCGATGATGGAGAGAAAATCATG
GCCACATCGTCACAGTGGCTTCAGTGTGCGGCCACGAAGGGATTCTTACCTCATCCCATA
TGTTCCAGCAAATTTGCCGCTGTGTGGCTTTCACAGAGGTCTGACATCAGAACTTCAGGCCT
GGGAAAACTGGTATCAAAACCTCATGTCTCTGCCAGTTTTTGTGAATACTGGGTTACCA
AAAATCAAGCACAAGATTATGGCCTGTATTGGAGACAGATGAAGTCGTAAGAAGTCTGATA
GATGAATACTTACCAATAAGAAAATGATTTTTTGTTCCTCATGATCAATCAATCTTTCTGAG
ACTACAGAAGTTTTCTCTGAACGCGCTCAGCGATTTTAAATCGTATGCAGAAATTTCAAT
TTGAAGCAGTGGTTGGCCACAAAATCAAATGAAATGAATAATAAGCTCAGCCAGACGATG
TATGCATGATAATGATATGAATAGTTTCGAATCAATGCTGCAAGCTTTATTTACATTTTT
TCAGTCTTGATAATATTA AAAACATTGGTTTGGCACTAGCAGCAGTCAAACGAACAAGATTA
ATTACCTGTCTTCTGTGTTCTCAAGAATATTTACGTAGTTTTTTCATAGGTCTGTTTTTCTCT
TCATGCCCTCTTAAAACTTCTGTGCTTACATAAACATACTTAAAAGGTTTTCTTTAAGATAT
TTTTTTTTTCCATTTAAAGGTGGACAAAAGCTACCTCCCTAAAAGTAAATACAAGAGAACT
TATTTACACAGGGAAGGTTTAAAGACTGTTCAAGTAGCATTTCAATCTGTAGCCATGCCACAG
AATATCAACAAGAACACAGAAATGAGTGCACAGCTAAGAGATCAAGTTTCAGCAGGCAGCTTT
ATCTCAACCTGGACATATTTTTAAGATTCAAGCATTTGAAAGATTTCCCTAGCCTCTTCTTTT
TCATTAGCCCAAAACGGTGCAACTCTATTCTGGACTTTATTACTTGATTCTGTCTTCTGTAT
AACTCTGAAGTCCACCAAAAGTGGACCTCTATATTTCTCTCCCTTTTATAGTCTTATAAGA
TACATTATGAAAGGTGACCGACTCTATTTTAAATCTCAGAATTTTAAAGTTCTAGCCCCATGA
TAACCTTTTTCTTTGTAAATTTATGCTTTCATATATCCTTGGTCCAGAGATGTTTAGACAAT
TTTAGGCTCAAAAATTAAGACTAACACAGGAAAAGGAAGCTGTACTGGCTATTACATAAGAAA
CAATGGACCAACAGAGAAGAA

FIGURE 192

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56409

<subunit 1 of 1, 300 aa, 1 stop

<MW: 33655, pI: 9.31, NX(S/T): 1

MNIILEILLLLITIIYSYLESLVKFFIPQRRKSVAGEIVLITGAGHGIGRQTTFYFAKRQSI
LVLWDINKRGVEETAAECKRLGVTAHAYVVDCSNREEIYRSLNQVKKEVGDTVIVVNNAGTV
YPADLLSTKDEEITKTFEVNILGHFWITKALLPSMMERNHGHIVTVASVCGHEGIPYLIPIYC
SSKFAAVGFHRGLTSELQALGKTGIKTSCLCPVFNVTGFTKNPSTRLWPVLETDEVVRSLID
GILTNNKMIFVPSYINIFLRLQKFLPERASAILNRMQNIQFEAVVGHKIKMK

Important features:

Signal peptide:

amino acids 1-19

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 30-33 and 58-61

Short-chain alcohol dehydrogenase family protein

amino acids 165-202, 37-49, 112-122 and 210-219

FIGURE 193

CGGCGGCGGCTGCGGGCGCGAGGTGAGGGGCGCGAGGTGAGGGGCGCGAGGTTCCAGCAGG
ATGCCCCGGCTCTGCAAGAGCTGAAGTGAGAGGCCGGAGAGGGCCCAGCCCCCGGGGG
AGGATGACCAAGGCCCGGCTGTTCGGCTGTGGTGGTGTGGGGTCGGTGTTTCATGATCCT
GCTGATCATCGTGTACTGGGACAGCGCAGGCGCCGCGCACTTCTACTTGCACACGTCCTTCT
CTAGGCCGCACACGCGGGCGCCGCTGCCACGCCCGGGCCGGACAGGGACAGGAGCTCACG
GCCGACTCCGATGTGACGAGTTTCTGGACAAGTTTCTCAGTGCTGGCGTGAAGCAGAGCGA
CCTTCCAGAAAGGAGACGAGCAGCGCCTGCGCCGGGGAGCATGGAGGAGAGCGTGAGAG
GCTACGACTGGTCCCCGCGCAGCCCGGCGCAGCCCAGACCAGGGCCGCGCAGCAGGCGGAG
CGGAGGAGCGTGCTGCGGGGCTTCTGCGCCAACCTCCAGCCTGGCCTTCCCCACCAAGGAGCG
CGCATTGACGACATCCCCAACTCGGAGCTGAGCCACCTGATCGTGGACGACCGGCACGGGG
CCATCTACTGCTACGTGCCCCAAGGTGGCCTGCACCAACTGGAAGCGCGTGATGATCGTGCTG
AGCGGAAGCCTGTGTCACCGCGGTGCGCCCTACCGCGACCCGCTGCGCATCCCGCGCGAGCA
CGTGACAACGCCAGCGCGCACCTGACCTTCAACAAGTTCTGCGCGCCGCTACGGGAAGCTCT
CCCGCCACCTCATGAAGGTCAAGCTCAAGAAGTACACCAAGTTCCCTTTCGTGCGCGACCCC
TTCGTGCGCCTGATCTCGGCCTTCCGAGCAAGTTGAGCTGGAGAACGAGGAGTTTCTACCG
CAAGTTCGCCGTGCCCATGTGCGGCTGTACGCCAACACACCAGCCTGCCCGCTCGGCGC
GCGAGGCCTTCCGCGCTGGCCTCAAGGTGTCCTTCGCCAACTTCATCCAGTACCTGCTGGAC
CCGCACACGGAGAAGCTGGCGCCCTTCAACGAGCACTGGCGGCAGGTGTACCGCCTCTGCCA
CCCGTGCCAGATCGACTACGACTTCGTGGGGAAGCTGGAGACTCTGGACGAGGACGCCGCGC
AGCTGCTGCACTACTCCAGGTGGACCCGCGAGCTCCGCTTCCCCCGAGCTACCGGAACAGG
ACCGCCAGCAGCTGGGAGGAGGACTGGTTCGCCAAGATCCCCCTGGCCTGGAGGCAGCAGCT
GTATAAACTCTACGAGGCCGACTTTGTTCTCTTCGGCTACCCCAAGCCCGAAAACCTCCTCC
GAGACTGAAAGCTTTCGCGTTGCTTTTTCTCGCGTGCTGGAACCTGACGCACGCGCACTCC
AGTTTTTTTATAGCTACGATTTTGCAATCTGGGCTTCTTGTTCACTCCACTGCCTCTATCC
ATTGAGTACTGTATCGATAATTGTTTTTTAAGATTAATATATTTTCAGGTATTTAATACGA

FIGURE 194

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56112

<subunit 1 of 1, 414 aa, 1 stop

<MW: 48414, pI: 9.54, NX(S/T): 4

MTKARLFRLWLVLGSVFMILLIIVYWDSAGAAHFYLHTSFSPHTGPPLPTPGPDRDRELT
DSDVDEFLLDKFLSAGVQSDLPKETEQPPAPGSMEESSVRGYDWSPRDARRSPDQGRQQAER
RSVLRGFCANSSLAFPTKERAFDDIPNSELSHLIVDDRHGAIYCYVPKVACTNWKRVMIVL
GSLLRGAPYRDPLRIPREHVHNASAHLTFNKFWRRYGKLSRHLMKVKLKKYTKFLFVRDPF
VRLISAFRSKFELNEEFYRKFAVPMRLRYANHTSLPASAREAFRAGLKVSFANFIQYLLDP
HTEKLAPFNEHWRQVYRLCHPCQIDYDFVGKLETLDEDAQLQLQVDRQLRFPPSYRNR
ASSWEEDWFAKIPLAWRQQLYKLYEADFVLFGYPKPENLLRD

Important features:

Signal peptide:

amino acids 1-31

N-glycosylation sites.

amino acids 134-137, 209-212, 280-283 and 370-373

TNFR/NGFR family cysteine-rich region protein

amino acids 329-332

THE UNIVERSITY OF CHICAGO

TGGGGCCAGAATTCCGCCACGAGGCCGACGAGGGCGACGGCCTCACGGGGCTTTGGAGGTGA
AAGAGGCCCCAGAGTAGAGAGAGAGAGACCGACGTACACGGGATGGCTACGGGAACCGGCT
ATGCCGGGAAGGTGGTGGTCTGTACCGGGGGCGGGCGCGGCATCGGAGCTGGGATCGTGCGC
GCCTTCTGTGAACAGCGGGGCCCGAGTGGTTATCTGCGACAAGGATGAGTCTGGGGGCCGGGC
CCTGGAGCAGGAGCTCCCTGGAGCTGTCTTTATCTCTGTGATGTGACTCAGGAAGATGATG
TGAAGACCTTGGTTTCTGAGACCATCCGCCGATTGGCCGCCTGGATTGTGTTGTCAACAAC
GCTGGCCACCACCCACCCCCACAGAGCCTGAGGAGACCTCTGCCCAGGGATTCCGCCAGCT
GCTGGAGCTGAACCTACTGGGGACGTACACCTTGACCAAGCTCGCCCTCCCTACCTGCGGA
AGAGTCAAGGGAATGTATCAACATCTCCAGCCTGGTGGGGCAATCGGCCAGGCCCAGGCA
GTTCCCTATGTGGCCACCAAGGGGGCAGTAAACAGCATTGACCAAAGCTTTGGCCCTGGATGA
AAGTCCATATGGTGTCCGAGTCAACTGTATCTCCCCAGGAAACATCTGGACCCCGCTGTGGG
AGGAGCTTGGCAGCCTTAATGCCAGACCTTAGGGCCACAATCCGAGAGGGCATGCTGGCCCGAG
CCACTGGGCCCGATGGGCCAGCCCGCTGAGGTGCGGGCTGCGGCAGTGTTCTTGCCCTCCGA
AGCCAACCTTCTGCACGGGCATTGAACTGCTCGTGACGGGGGTGTCAGAGCTGGGGTACGGGT
GCAAGGCCAGTTCGGAGCACCCCGTGAGCGCCCCCGATATCCCTTCCTGAATTCTCTCATT
CTACTTGGGGCCCCCTTCTTAGGACTCTCCACCCCAAACCTCCAACCTGTATCAGATGCAGC
CCCCAAGCCCTTAGACTCTAAGCCAGTTAGCAAGGTGCCGGGTACCCCTGCAGGTTCCTCAT
AAAAACGATTTCCAGCC

FIGURE 196

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56045

<subunit 1 of 1, 270 aa, 1 stop

<MW: 28317, pI: 6.00, NX(S/T): 1

MATGTRYAGKVVVVTGGGRGIGAGIVRAFNNSGARVVICDKDESGGRALEQELPGAVFILCD
VTQEDDVKTIVSETIRRFGRLCDVNVNAGHHPPQRPETSAQGFRQLLELNLLGTYTLTKL
ALPYLRKSQGNVINISSLVGAIGQAQAVPYVATKGAVTAMTKALALDESPYGVRVNCISPGN
IWTPLWEELAALMPDPRATIREGMLAQPLGRMGQPAEVGAAAVFLASEANFCTGIELLVTGG
AELGYGCKASRSTPVDAPDIPS

Important features:

N-glycosylation site.

amino acids 138-141

Short-chain alcohol dehydrogenase family protein

amino acids 10-22, 81-91, 134-171 and 176-185

FIGURE 197

AGGCGGGCAGCAGCTGCAGGCTGACCTTGCAGCTTGGCGGA**ATG**ACTGGCCTCACAACCTG
CTGTTTCTTTCTTACCATTTCATCTTCTGGGGCTGGGCCAGCCCAGGAGCCCCAAAAGCAA
GAGGAAGGGGCAAGGGCGGCCTGGGCCCCCTGGCCCCCTGACAGGTGCCACTGGACC
TGGTGTACGGATGAAACCGTATGCCCGCATGGAGGAGTATGAGAGGAACATCGAGGAGATG
GTGGCCAGCTGAGGAACAGCTCAGAGCTGGCCAGAGAAAGTGTGAGGTCAACTTGCAGCT
GTGGATGTCCAACAAGAGGAGCCTGTCTCCCTGGGGCTACAGCATCAACCACGACCCAGCC
GTATCCCCGTGGACCTGCCGGAGGCACGGTGCCCTGTGTCTGGGCTGTGTGAACCCCTCACC
ATGCAGGAGGACCGCAGCATGGTGAGCGTGCCGGTGTTAGCCAGGTTCTGTGCGCCGCCG
CCTCTGCCCGCCACCGCCCCGCACAGGGCCTTGCCGCCAGCGCGCAGTCATGGAGACCATCG
CTGTGGGCTGCACCTGCATCTTCT**TGA**ATCACCTGGCCAGAAGCCAGGCCAGCAGCCCGAGA
CCATCCTCCTTGACACCTTTGTGCCAAGAAAGGCCTATGAAAAGTAAACACTGACTTTTGAAA
GCAAG

1000
900
800
700
600
500
400
300
200
100
0

FIGURE 198

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA59294

<subunit 1 of 1, 180 aa, 1 stop

<MW: 20437, pI: 9.58, NX(S/T): 1

MDWPHNLLFLLTISIFLGLGQPRSPKSKRKGQGRPGPLAPGPHQVPLDLVSRMKPYARMEEY
ERNIEEMVAQLRNSSELAQRKCEVNLQLWMSNKRSLSPWGYSINHDPRIPVDLPEARCLCL
GCVNPFTMQEDRSMVSVFVSQVPVRRRLCPPPPRTGPCRQRAVMETIAVGCTCIF

Important features:

Signal peptide:

amino acids 1-20

N-glycosylation site.

amino acids 75-78

Homologous region to IL-17

amino acids 96-180.

FIGURE 199

GCGCCGCCAGGCGTAGGCGGGGTGGCCCTTGCGTCTCCCGCTTCTCTTGAAAAACCCGGCGGG
CGAGCGAGGCTGCGGGCCGGCCGCTGCCCTTCCCCACACTCCCGCCGAGAGGCCTCGCTCG
GCGCCCAACATGCGCGGGTGGGCGCTGCGGCCCGCAGCTAACGGCGCTCCTGGCCGCTTGGAT
CGCGGCTGTGGGCGGACGGCAGGCCCCGAGGAGGCCGCGCTGCCGCCGAGCAGAGCCGGG
TCCAGCCCCATGACCCGCTCCAACCTGGACGCTGGTGATGGAGGGCGAGTGGATGCTGGAATTT
TACGCCCCATGGTTCTCATCTCTGCCAGCAGACTGATTGAGAATGGGAGGCTTTTGGCAAGAA
TGGTGAARATCTTCAGATCAGTGTGGGGAAGGTAGATGTCATTCAAGAACCAGGTTTGTAGT
GCCGCTTCTTTGTCAACCACTCTCCAGCATTTTTCATGCAAAAGGATGGGATATTCCGCCGT
TATCGTGGCCAGGAATCTTCAAGACCTGCAGAATTATATCTTAGAGAAGAAATGGCAATC
AGTCGAGCCCTCTGACTGGCTGGAAATCCCCAGCTTCTCTAACGATGTCTGGAATGGCTGGTC
TTTTTAGCATCTCTGGCAAGATATGGCATCTTCACAACTATTTACAGTGA CTCTTGGAAAT
CCTGCTTGGTGTCTTATGTGTTTTTCGTCATAGCCACCTTGGTTTTTGGCCCTTTTTATGGG
TCTGGCTTGGTGGTAAATATCAGAATGTTTCTATGTGCCACTTCCAAGGCATTTATCTGAGC
GTTCTGAGCAGAATCCGAGATCAGAGGAGGCTCATAGAGCTGAACAGTTGCAGGATGCGGAG
GAGGAAAAAGATGATTCAAATGAAGAAGAAAAACAAGACAGCCTTGTAGATGATGAAGAAGA
GAAAGAAGATCTTGGCGATGAGGATGAAGCAGAGGAAGAAGAGGAGGAGGACAACCTTGGCTG
CTGGTGTGGATGAGGAGAGAAGTGAGGCCAATGATCAGGGGCCCCAGGAGAGGACCGGTGTG
ACCCGGGAGGAAGTAGAGCCTGAGGAGGCTGAAGAAGGCATCTCTGAGCAACCTTGCCAGC
TGACACAGAGGTGGTGAAGACTCCTTGAGGCAGCGTAAAGTCAGCATGCTGACAAGGGAC
TGTAGATTTAATGATGCGTTTTCAAGAATACACACAAAAACAATATGTCAGCTTCCCTTTGG
CCTGCAGTTTGTACCAAAATCCTTAATTTTTTCTGAATGAGCAAGCTTCTCTTAAAGATGCT
CTCTAGTCATTTGGTCTCATGGCAGTAAGCCCTCATGTATACTAAGGAGAGCTTCCAGGTGT
GACAATCAGGATATAGAAAAACAACGTAAGTGTGGGATCTGTTTGGAGACTGGGATGGGAA
CAAGTTCATTTACTTAGGGTCAGAGAGTCTCGACCAGAGGAGGCCATTCCAGTCCCTAATC
AGCACCTTCAGAGACAAGGCTGCAGGCCCTGTGAATGAAGCCCAAGCAGGAGCCTTGGCT
CCTGAGCATCCCCAAAGTGTAACGTAGAAGCCTTGATCCTTTCTTGTGTAAGTATTAT
TTTTGTCAAATTCAGGAAACATCAGGCACCCAGTGTCATGAAAAATCTTTACAGCTAGAA
ATTGAAAGGGCCTTGGGTATAGAGAGCAGCTCAGAAGTCATCCCCAGCCCTCTGAATCTCTG
TGCTATGTTTTATTCTTACCTTTAATTTTTTCCAGCATTTCACCATGGGCATTACAGGCTCT
CCACACTCTTCATATTATCTCTTGGTCAGAGGACTCCAATAACAGCCAGGTTTACATGAAC
TGTGTTTTGTTTCTGACCTAAGGGGTTTGAATAATCAGTAACCAACCCCTGAAGCTGT
GACTGCCAAACATCTCAAATGAAATGTTGTGGCCATCAGAGACTCAAAGGAAGTAAGGATT
TTACAAGACAGATTAAAAAAATTTGTTTTGTCCAAATATAGTTGTTGTTGATTTTTTTTT
AAGTTTTCTAAGCAATATTTTTCAAGCCAGAAGTCCTTAAGTCTTGCAGCTACAAGGTAGT
CTTGTGAAGAAAAGTTGAATACTGTTTTGTTTTCATCTCAAGGGGTTCCCTGGGCTTGAAC
TACTTTTAATAAATAACAAAAACCCTCTGATTTTCTTCAAGTATGCTGTTTGGTGAAA
GAATTAATGAACTCCAGTACCTGAAAGTGAAAGATTGATTTTTGTTCCATCTCTGTGAATC
TTCCAAGAATATATATCTTTGTAATCTCTCAATACTCAATCTACTGTAAGTACCCAGGGAG
GCTAATTTCTTT

FIGURE 200

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56433

<subunit 1 of 1, 349 aa, 1 stop

<MW: 38952, pI: 4.34, NX(S/T): 1

MAGGRCGPQLTALLAAWIAAATAAGPEEAALPPEQSRVQPMTASNWTLVMEGEWMLKFYAP
WCPSCQQTDSWEAEFAKNGETLQISVGKVDVIQEPGLSGRFFVTTLPAPFFHAKDGIFFRYRG
PGIFEDLQNYILEKKWQSVEPLTGWKSPASLTMSGMAGLFSISGKIWHLHNYFTVTLGIPAW
CSYVFFVIATLVFGLFMGLVLVVISSECFYVPLPRHLSESRSEQNRRSEEAHRAEQLQDAEEEEK
DDSNEEENKDSLVDDEEEKEDLGDEDAEEEEEEEDNLAAGVDEERSEANDQGPPGEDGVTRE
EVEPEEAEEGISEQPCPADTEVVEDSLRQRKSHADKGL

Important features:

Signal peptide:

amino acids 1-22

Transmembrane domain:

amino acids 191-211

N-glycosylation site.

amino acids 46-49

Thioredoxin family proteins. (homologous region to disulfide isomerase)

amino acids 56-72

Flavodoxin proteins

amino acids 173-187

[illegible]

ATCTGGTTGAAC TACTTAAGCTTAATTTGTTAAACTCGGGTAAGTACCTAGCCACATGATT
GCATCCAGATATCTCTTTGTCCACAGACAGCTCATCTCAGGGGCAGAAAGAAAGCTCC
CAATCGCTATATCTATTACAGGGGCTCTCAGAAACAATTGGAATATCATCTCGATTATTAGAAAA
TTGGATGAAGATGGATATATCAATTAACCTACGCTCATCTCAAGCAATACAGGATAGCTGT
TGTTTCAGAGAAAGGATCGTGTGCTGCATCTCTCTCTGGCGCTCATCTGCTGTAATTTTGG
GAATCCTATGTTCTGGTAATACCTGGTGATAGCTGTGGTCTCTGGGTACCATGGGGGTTCTTTCC
AGCCCTTGCTCTCTTAATTTGGATTATATAGAAAGAGCTGTATCTATTCCAGCATGTGCTACT
AAATTCCTGGGATGGAAGTAAAGACAATGGCTGGCACTGGGCTCTAATCTCTCAAGATAG
ACAGCTCAATGAATTTGGGATTATAGTAAACAAGTGTCTTCCCAACCTGATAATTCATT
TGCTAGGCGTTTCTCGCCCCAGCATGAGTATACCATGCTCTGGGAGGATGATCAACATT
CTCTTCTAACTTTTTCAGATCAGAAGCAAGCTACCCCAAGAAACCCATCTCCAATTTGTG
TAGGATTACACGTGTCAGTCACTTATGACCAACTGTGTAGTGTGCCCTCATATAGAATTTGT
GAGAGAAGTTTTCATGTAAAGGAGGAGGTGGAGAGAGAGAGAGAATATGTGAGGTAGTA
AGGAGAGACAGAAAACGAACAGAAAGAGATTAACAGCTGAGGTACAGGATAAATGACAGAAATG
TTTAGAGAGCTTGGCCAACTGTAATCTTAAACGAATTTGAAGGAGAGGCTGTGATTTC
GTATTTGTCGACCTACAGGTAGGCTAGTATTATTTTCTAGTTAGTAGATCCCTAGACATGG
AATCAGGGCAGCGCAAGCTTGAGTTTATTTTATTTATTTATTTTGTAGATAGGTTCT
CAGTTTGTATACCCAGGCTGGAGTGCAGTGGCAACTCATGCACTGCAGCATCTCTCGC
CTCAGCCGCTCAAGTAGCTGGGACTGAGTGCATGACCCATGCGCAGGCTAATTTTGTGT
TTTTTTGTAGACATGGGTTTGGCCATGTGTGACCAAGCTGCTCTCAACTCTGGGCTTAAG
TGATCTGCCCGCCTTGGCCTCCCAAAGTGTGGGATTACAGATGTGAGCCACCACACCTGGC
CCCAAGCTTGAATTTTCACTTCGCCATTGACTGGCATTACCTTGGGTAAAGCCATAAGCGA
ATCTTAATTTCTGGCTCTACAGAGTTGTTTCATGCTCAACAATGCCATTGAAGTGCAGGCT
GTGTTGCCACGATTAGCCCTCAACTCTATGACAGATATACGTTATGAACCTGAGGCTGAAAT
ATATTTCTGAATAGCTAAATAGGAATATGGAGAAAAATCTCACCACTACGACGCAATTTT
ATTATTTTTCATCAGTATGATCATAAATATGATTATCATCTTAGTAAAGACAGGAATCTTA
CTTTTCTTTTCAATTAATAGCTCAGAGATACATCTGCCATATCTCTAATAGAACTCTT
TTTTTTTTTTTTTTTTGAGACAGAGTTTCGCTCTGTGTGCCAGGCTGGAGTGCACCGG
CAGATCTCGGCTCAGCCGAACCTCGGCCCGCTGGGTTCAAGCAATTTCTCTGCTCCAGCT
CCCAAGTCTGGGATTACAGTCAGGCACCCACACCCGGCTAATTTGTATTTTGTATTTT
AGAGACAGGGTTTCTCCATGTCTCGGTGAGGTAGTCCGGAATCCTGACCTCAAGTGAATCTG
CTGCTCTGGCCTCCCAAGTGTGGGATTACAGCGGTGAGGCTGACCTCCAGCCAGATAGATCT
TGTATAATATGTAATTTAGGGAAACTGCTCTCATAGGAAAGTTTCTGCTTTTAAATACA
AAAAATACATAAAATACATAAAATCTGATGATGAATATAAAAGAGTAACCACTCAATTGGA
ACAGATATTAACAAATTTGGAATATTTTATAGTTTGTGATGTACTGTTTACAATTTT
ACCATTTTTTTCAGTAATTACTGTAAATATGTTATTTGGAATGAACATATATTTCTCATG
TCGTGATTTGTCTATTTTTCATCTTTCCCACTGGTGTCTATTTTATTTCCATGTGATA
TTTTCTGATTACTAGGGAGGCATTACAGCTCTCTAATGTGATTAAATATGTGAAAGAAAT
TGTAACAAATTTTACTAATATTTACGTTTAAAAATGGATGATTTATGTATGTGGATTTCAT
TTCATAAAAAAAAACTTATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 202

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53912

<subunit 1 of 1, 201 aa, 1 stop

<MW: 22563, pI: 4.87, NX(S/T): 1

MEYHPDLENLDEDGYTQLHFDSQSNTRIAVVSEKGS CAASPPWRLIAVILGILCLVILVIAV
VLGTMGVLSSPCPPNWIIEKSCYLFMSLSNSWDGSKRQCWLGSNLLKIDSSNELGFIKQ
VSSQPDNSFWIGLSRPQTEVPWLWEDGSTFSSNLFQIRTTATQENPSPNCVWIHVSVIYDQL
CSVPSYSICEKKFSM

Important features:

Type II transmembrane domain:

amino acids 45-65

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 197-200

N-myristoylation sites.

amino acids 35-40 and 151-156

Homologous region to LDL receptor

amino acids 34-67 and 70-200.

FIGURE 203

GGAAGGGGAGGAGCAGGCCACACAGGCCAGGCCGGTGAGGGACCTGCCAGACCTGGAGGGTCTCGCTCTGTCA
 CACAGGCTGGAGTGCAGTGGTGTGATCTTGGCTCATCGTAACCTCCACCTCCCGGGTTCAGTGATTCTCATGCG
 TCAGCTCCCGAGTAGCTGGGATTACAGTGGTGACTTCCAAGATGACTCCGTCCGAGGAGAAATGACTCCCCAG
 TCGCTGCTGCACAGACACTGTCTTCTGCTGAGTCTGCTCTTCTGGTCCAAGGTGCCCAAGCGGAGGGGCCACAGG
 GAAGACTTTCCTCTCTGCAGCAGCGAAACACAGACACAGGAGCAGCCTCCACTCAAAACCCACACAGCACTCG
 CGCATCTCCATCGAGAACTCCGAAGAGGCCCTCACAGTCCATGCCCTTTCCTGTCAGCCCACTCGCTTCCCGA
 TCCCTTCCCTGACCCCGGGCCCTTACCACTTCTGCTCTACTGGAACCGACATGTCTGGGAGATTACATCTTCTC
 TATGGCAAGCGTGACTTCTTGTGTGAGTGACAAAGCCTCTAGCCTCTCTGTCTTCCAGCACAGGAGGAGGAGCCTG
 CCTCAGGGGCCCGCGCTGTAGCACTTCTGTGCACCTCTCGGTGGAGCCCTCAGAACATCAGCCTGCCCATGTGCC
 GCCAGCTTCACTTCTCTTCCACAGTCTCCCCACAGGCCCGCTCACAATGGCTCGGTGGACATGTGCGAGCTC
 AAAAGGGACCTCCAGCTGTCTCAGCCAGTCTCTGAAGCATCCCCAGAGGCCCTCAAGGAGGCCCTCGGCTGCCCGC
 GCCAGCCACGACTTGCAGAGCCTGGAGTCGAAACTGACCTCTGTGAGATTCTATGGGGACATGGTGTCTCTCGAG
 GAGGACCGGATCAACGCCACGGTGTGGAAGCTCCAGGCCACAGCCGGCTCCAGGACCTGCACATCCACTCCCGG
 CAGGAGGAGGAGCAGAGCGAGATCATGGAGTACTCGTGCTGTGCTCGCTCGAACACTTCTCCAGGAGCGAAAGGC
 CGGAGCGGGGAGGCTGAGAGAGACTCTCTCGTGGTGGACTTCAGCAGCCAGGCCCTGTCCAGGACAAAGATTCC
 AGCCAAGTCTCGGTGAGAAAGTCTTGGGATTGTGGTACAGAACACCAAAGTAGGCCAACCTCAGCGAGCCCGTG
 GTGCTCACTTCTCCAGCACCGACTCAGCCGGAAGAAATGTGACTCTGCAATGTGTGGTCTTGGGTTGAAGACCCCA
 TTTGAGCCCGCGGGCATTTGGAGCAGTGTGGGTGTGAGACCGTCCAGGAGGACAAACCCCACTCGCTCTGTG
 AACCATCTGACCTACTTTGCAGTGTGATGGTCTCTCGTGGAGGTGGAGCGCGTGCACAAAGCACTACTGCTCAGG
 CTCTCTCTCTGAGTGGGCTGTGCTGTCTGCGCCTGGCTGCTGTCAACATTGCTGCTCGCTCTGCTCTCAGG
 GTGCGCCTCGCTGTCAGGAGGAAAGCTCGGGACTACACCATCAAGGTGCACATGAACCTGTCTGCGGCCGTCTTC
 CTGCTGGACAGTCTCTGCTCAGCGAGCGGTGGCCCTGACAGGCTCTGAGGCTGGCTGCGGAGCCAGTGGC
 ATCTTCTGCACTTCTCCTGTCTCCTGCTCTTCTGGATGGGCTCCGAGGGGTACAACCTCTACCGACTCTGT
 GTGGAGGTCTTTGGCACTTGTCTCTGGCTACTACTCAAGCTGAGCCCATCTGGCTGGGCTTCCCCATCTTT
 CTGCTGACCTGTGGGCCCTGTGGATGTGGACAACTATGGCCCATCATCTTGGCTGTGCTATGAGTCCAGAG
 GGGCTCATCTACCTCTCCATGTGTGATTCGGGACTCCCTGTCTAGCTACATCAACACCTCTGGCCCTCTTCAGC
 CTGGTGTCTTCTGTTCAACATGCGCTGTCTAGCCACCATGGTGTGTCAGATCTCTGCGGCTTCCAGCCAA
 AAGTGGTCAATGTGCTGACTGCTGTGCTGGCTCAGCCTGGTCTTGGCCTGCCCTGGGCTTGATCTTCTCTCC
 TTTGCTTCTGCGACCTTCCAGCTGTGTCTCTCTACCTTTTCAGCATCATCAACCTCTTCCAAGGCTTCTCATC
 TTCACTGTGATCTGGTCCATGCGGCTCAGGCCCGGGGTGGCCCTCCCTCTGAAGAGCACTCAGACAGCGCC
 AGGCTCCCCATCAGCTCGGGCAGCACTCTGCCAGCCGATCTAGGCTCTGAGCCCATCTGCTGATGAAG
 CAGAGATGCGGCTCTGTGTCAGACTGCTGTGGGCCCGGAGCCAGGCCAGGCCAGGCGCAGCTCAGCCGACACT
 TTTGAAAGCCCAACGACCTGGAGAGATGGGCCGTGGCCATGGTGGAGCACTCCCGGGCTGGGCTTTGAATTG
 GCTTCTGGGACTACTCGGCTCTCACTCAGCTCCACCGGACTCAGAAGTGGCGGCCATGTGCTGCTAGGGTACTG
 TCCCAACATCTGTCGCAACCCAGCTGGAGGCTGGTCTCTCTTACAACCTGGGGCCAGCCCTCATTTGCTGGG
 GGGCAGGCCCTTGGATCTTGGAGGTCTGGACATCTCTAAATCTGTGCCCTGCTGGGACAGAAATGTGGCTCA
 GTTGTCTGTCTCTGTGTGTCACCTGAGGCCACTCTGCATCCTGTGTCATTTTAACTCAGGTGGCCACCCAGG
 GAGTGTGGGCCCGAGGCCAGACTTCAGGCCAGAGCCCTGGCGGAGGAGGCCCTTTGCCAGGAGCACAGCAGC
 CGCTCGCTACTCTCTGAGCCAGGCCCTCTCTCTCTCAGCCCCCAGCTCTCCCTCCATCTTCTCTGGGGTTC
 TCTCTCTCTCCAGGGCTCTCTGTCTCTCTGTTCAAGCTGGGGGTCCCCGATTCGAATGCTGTTTGTGGGGA
 GTGGTTTCCAGGAGCTGCTGTGTGTCTGTCTGTAATGTTTGTCTACTGCAACAGCTCGGCTGCCCTGAGCACA
 GGCTCGGTACCGATGCGTGGCTGGCTAGGTCCCTCTGTCCATCTGGGCCCTTGTATGAGCTGCAATGCCCTTG
 CTACCTCCAGCAAGACACCGCTCAGAGGGGCCCTCAGCCTCTCTGAGGCCCTCTGTGGCAAGAACTGTGGA
 CCATGCCAGTCCGCTGTGTTTCCATCCACCACTCCAAGGACTGAGACTGACTCTCTGTGTGACACTGGCCTTA
 GAGCCTGACATCTCTCAAGAGGTTCTCTCAAGCCCCAAATAGCTCCAGGCCCTCGGCCGCCCATCATGTGT
 TAACTCTCTCAACAAACACACAGCGGTAGATTGCTGGCCTGTGTAGTGTGTAGGACACAGATGACCCAGCTG
 GCTACTCCTCTGCCCAACTCAGTCTGGTATGTGAGGCGTGGTGAAGCAAGAACTCTGGGACTACAGGACA
 GGGAGCCATCATCTCTGCTCGGGAACTCTGGAAGACTTCTCTGAGGAGTCAGCGTTCATCTTGACCTTGAAGAT
 GGGAGAGTGTCTTTTATACCAATCTTTTGTCTTTGATATTAAAGAGAGTACATGTTCACTTGTAGAGA
 ATTTGGAACCTGTAGAAGAGAATCAAGAAGAAATAAAAATCAGTGTGTTAATCGCTAGCAAAAAA
 AA

3007194.102443

FIGURE 204

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50921

<subunit 1 of 1, 693 aa, 1 stop

<MW: 77738, pI: 8.87, NX(S/T): 7

MT PQSLLQTTLFLLSLLFLVQGAHGRGHREDFRFCSQRNQTHRSSLHYKPTPDLRISTENSE
EALT VHAPFPAAHPASRSFPDPRGLYHFCLYWNRHAGRLHLLYGKRDFLLSDKASSLLCFQH
QEE SLAQGPPLLATS VTSWVSPQNI SLPSAASF TFSFHSPPHTAAHNASVDMCELKRDLQLL
SQFLKHPQKASRRPSAAPASQQLQSLESKLT SVRFMGDMVSFEEDRINATVWKLQPTAGLQD
LHIHSRQEEEQSEIMEYSVLLPRTL FQRTKGRSGEAEKRLLLVDFSSQALFQDKNSSQVLGE
KVLGIVVQNTKANL TE PVVLT FQHQLQPKNVTLCVFWVEDPTLSSPGHWSSAGCETVRRE
TQTSCFCFNHLTYFAVL MVSSPEVD AVHKHYLSLLSYVGCVV SALACLV TIAAYLC SRVPLPC
RRKPRDYTIKVH MNLLLA VFLDTS FLLSEPVALTGSEAGCRASATFLHFSLLTCLSWMGLE
GYNLYRLVVEVFGTYVPGYLLKLSAMGWGFPIFLVTLVALVDVDNYGPIILAVHRTPEGVIY
PSMCWIRDSLVS YITNLGLFSLVFLFN MAMLATMVVQILRLRPHTQKWSHVLTLGLSLVLG
LPWALIFFSFSAGTFQLVVLVLYFSIITSFQGFLIFIWYWSMRLQARGGPSPLKSNSDSARLP
ISSGSTSSSRI

Important features:

Signal peptide:

amino acids 1-25

Putative transmembrane domains:

amino acids 382-398, 402-420, 445-468, 473-491, 519-537, 568-590
and 634-657

Microbodies C-terminal targeting signal.

amino acids 691-693

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 198-201 and 370-373

N-glycosylation sites.

amino acids 39-42, 148-151, 171-174, 234-237, 303-306, 324-327
and 341-344

G-protein coupled receptors family 2 proteins

amino acids 475-504

[illegible]

TGCGCTGGCGCTGCCTTGTCACAACATGCCGCTTACTCTGCTTCCAGGTTGCCCTGCCTTGCAGAG
GGAAANNCNTCGGGACTACACNCNTCAAGTGACATGAACCTGCTGCTGGCCGCTCTTCTCTGCTG
GACACGAGCTTCTCTGCTCAGCGNAGCCGCTGGCCCTGACAGGCTCTGAAGGCTGGCTGCCGA
GCCAGTGCCATCTTCTGCACTTCTCTCTGCTCACTGCCTTTCTTGATGGGCCCTCGAGGGG
TACAACCTCTACCGACTCTGTTGGTGAGGCTCTTTGGCACCTATGTCCCTGGCTACCTACTCAA
GCTGAGCGCCATGGGCTGGGGCTTCCCATCTTTCTGGTGACGCTGGTGGCCCTGGTGGATG
TGGCAACCTATGGCCCCATCATCTTGGCTGTGCATAGGACTCCAGAGGGCGTCACTACCTT
TCCATGTGCTGATCGGGACTCCCTGGTCAGCTACATCACCAACCTGGGCCCTTTCAGCCT
GGTGTCTCTGTTCAACATGG

FIGURE 207

MSLFGLLLLTALAGQRQGTQAESNLSSKFQFSSNKEQNGVQDPQHERIITVSTNGSIHSPR
FPHTYPRNTVLVWRLVAVEENVWIIQLTFDERFGLDEDPEDDICKYDFVEVEEPSDGTILGRWC
GSGTVPGKQISKGNQIRIRFVSDEYFPSEPGFCIHYNIVMPQFTEAVSPSVLPSPALPLDLL
NNAITAFSTLEDLIRYLEPERWQLDLEDLYRPTWQLLGKAFVFGRKSRVVDLNNLLTEEVRLY
SCTPRNFSVSIREELEKRTDTIFWPGCLLVKRCGGNCACCLHNCNECQCVPKVTKKYHEVLQ
LRPKTGVRGLHKS LTDVALEHHEECDVCRCGSTGG

Signal sequence:

amino acids 1-14

1017191102401

FIGURE 208

CCCATCTCAAGCTGATCTTGGCACCCTCTCATGTCTGTCTCTTCTCAACCAGACCTCTACATTCCATTTTGGGAAGA
 AGACTAAAAATGGTGTGTTTCCAAATGTGGACACTGAAGAGACAAATTCCTATCCTTTTAAACATAATCCTAATTTCC
 AAACCTCTTGGGGCTAGATGGTTTCTCAAAACTCTGCCCTGTGATGTCACTCTGGATGTTCCAAAGCAACCATGTG
 ATCTGGGACTGCAACAGACAAGCATTTGACAGAAATTCCTGGAGGTATCCCAAGAACCCACGAACTCACCTCT
 ACCATTAAACCATACAGACATCTCCCCAGCGTCCTTTCACAGACTGGGACCATCTGGTAGAGATCGATTTTCAGA
 TGCACCTGTGTAACCTATTCACCTGGGGTCAAAAAACAACATGTGCATCAAGAGGCTGCGAGATTAAACCCAGAAGC
 TTTAGCTGGACTACATTTATTTAAAAATCCCTTACCTGGATGGAAACAGGCTACTAGAGAGATACCGCAGGGCTCCCG
 CCTAGCTTACAGCTTCTCAGCCTTGGGCCAACAAATCTTTCCATCAGAAAAGAGATCTAACAGAACTGGCC
 AACATAGAAAATCTCTACCTGGGGCCAAAATCTGTTATTTATCGAAATCCTTGTATGTTTCATATTCAATAGAGAA
 GATGCCCTCTCAAACCTTGACAAAGTTAAAAAGTGCTCTCCCTGAAAGATAACAATGTCAACAGCGCTCCCTAGTGT
 TTGCCATCTACTTTAACAGAACTATATCTCTCAACAACATGATGCAAAAATCCAAGAGATGATTTTAAATAAC
 CTCAACCAATTAACAAATCTTGACCTTAAGTGGAAATTGCCCTCGTTGTATATAATGCCCATTTCCCTTGTGGCCG
 TGTAAAAATAATTTCTCCCTACAGATCCCTGTAAATGCTTTTGTATGCGCTGACAGAAATTAAGATTTTACGTCTA
 CACAGTAACTCTTTCAGCATGTGCCCCCAAGATGGTTTAAAGAACATCAACAAATCCAGGAACCTGAGATCTGCC
 CAAAATCTCTTGGCCAAAAGAAATGGGGATGCTAAATTTCTGCAATTTCTCCCCAGCCTCATCCAATTTGGATCTG
 TCTTTCAATTTTGAACCTTCAGGTCTATCGTGCATCTATGAATCTATCAACAAGCATTTTCTTCACTGAAAAGCCTG
 AAAATTTCTGGGATCAGAGGATATGTCTTTAAAGAGTTGAAAAGCTTAAACCTCTGCCATTCACATAATCTTCAA
 AATCTTGAAGTCTTGGATCTTGGCACTAACTTTATAAAAAATGCTAACTCAGCATGTTTAAACAAATTTAAAGA
 CTGAAAGTCAATAGATCTTTCAGTGAATAAAATATCACTTTCAGGAGATTCAAGTGAAGTTGGCTTCTGCTCAAAT
 GCCAGACTCTCTGTAGAAAGTTATGAACCCAGGCTCTGGAACAAATACATATTTTCAGATATGTAAGTATGCA
 AGGAGTTGACAGTTCAAAAAGAAAGAGGCTTCTTTCATGTCTGTAAATGAAGCTGCTACAAGATATGGGACAGCC
 TTGGATCTAAGTAAAAATAGTAAATTTTGTGCAAGTCTCTGATTTTCAACATCTTCTTCTTCAATGCTCCG
 AATCTGTGCAAGAACTCATTTAGCCAACTCTTAATGGCAGTGAATTCACAACTTTTACAGAGCTGAGATATTTG
 GACTCTCTCAACAAACCGCTGTGATTTTACTTCATCAACAGCAITTGAAGAGCTTCAACAACTGGAAGTCTGTGAT
 ATAAAGCAATAAGCAATTTTTCATCAAGAGGAATTAATCTATATGCTAAACCTTTACCAAGAACTTAAAGGTT
 CTGCAAAATGATGATGAACGCAATGACATCTCTTCTCCACAGCAGAGACCATGGAGATGTGATCTCTTAGA
 ACTCTGGAATTCAGAGAAATCTCCTAGATGTTTATGGAGAGAGGTTACACAGATACCTTACAATTTCAAG
 AATCTGTCTAAATTTAGAGGAATAGACATCTCTAAAAATTCCTTAAGTTTCTGTGCTCTGTGGAGTTTGTGATG
 ATGCCCTCAAATCTAAAGAACTCTCTTTGGCCAAAATGGGCTCAAATCTTTCAGTTGGAAGAACTCCAGTGT
 CTAAGAACTCTGGAACCTTGGACCTCAGCCCAACCAACTGACCACTGTCTCTGAGAGATTATCCAATCTGTCTC
 AGAAGCTTCAAGAACTGATCTCTTAAAGAAATAACAAATCAGGAGTCTGACGAAGTATTTCTACAAGATGCCCTT
 CAGTTGCGATATCTGGATCTCAGCTCAAATAAAATCCAGATGATCCAAAAGACCACTCTCCAGAAAATGTCTCT
 AACAACTCTGAAGATGTTGCTTTGCAATATAATCGGTTTCTGTGCACCTGTGATGCTGTGTGGTTTGTCTGTG
 GTTAAACATACGAGGATGACTATTCCTTACTCTGGCCACAGATGTGACTTGTGTGGGGCCAGGAGCACACAGGGC
 CAAAGTGTGATCTCCCTGGATCTGTACACCTGTGAGTTAGATCTGACTAAACCTGATTTCTGTTCTTCCATA
 TCTGTATCTCTTCTCTCATGTGGATGATGACAGCAAGTCACTCTATTTCTGGGATGTGTGGTATATTTACCAT
 TTTCTGAAGGCCAAGATAAAGGGGTATCAGCGCTTAATATCACAGACTGTGTTCTATGATCTTTTATTTGTGAT
 GACATTAAGACCCAGCTGTGACCGAGTGGGTTTGTGCTGAGCTGTGTGGCCAACTGGAAGACCCAGAGAGAAA
 CATTTTAAATTTATGTTCTCAGGAAAGGGACTGGTTACCAAGGGCAGCCAGTTCTGGAACAACTTCCAGAGCATA
 CAGCTTAGCAAAAAAGACAGTGTGTTGTGATGACAGACAAGTATGCAAGAAGCTGAAAATTTTAAAGATGCAATTTAC
 TTGTCCATCAGAGGCTCTAGGATGAAAAGTTGATGTGATATCTTGTATTTCTTGAAGAGCCCTTTCAAGAG
 TCCAAGTTCTCCAGCTCCGGAAGAGGCTCTGTGGGAGTCTGTCTCTGATGTGGCCAAACAAACCCGCAAGCTCAC
 CCATATCTTGGCAGTGTCTTAAAGAACGCCCTGGCCACAGACATCATGTGGCCATAGTCTCAGGTGTTCAGGAA
 ACGGCTAGACCTTCTTTGCAAAACACAACTGCTTAGTTTACAAAGAGAGGCTGGC

10017191-100191

FIGURE 209

MVFPMWTLKRQILILFNIIISKLLGARWFKTLPDVTLDVDPKNHIVDCTDKHLTEIPGG
IPTNTTNLTLTINHIPDISPASFHRLDHLVEIDFRCNCVPIPLGSKNNMCIKRLQIKPRSFS
GLTYLKSLEYLDGNQLLEIPQGLPPSLQLLSLEANNIFSIRKENLTELANIEILYLGQNCYYR
NPCYVYSYSIEKDAFNLNLTFLKVLSLKDNNTAVPTVLPSTLTLYLYNNMIAKIQEDDFNNL
NQLQILDLSGNCPRCYNAPFPAPCKNNSPLQIPVNAFDALTELKVLRLHSNSLQHVPPRWF
KNINKLQELDLSQNFLAKEIGDAKFLHFLPSLIQLDLSFNFLQVYRASMNLSQAFSSLSL
KILRIRGYVFKELKSFNLSPLHNLQNLEVLDLGTNFIKIANLSMFKQFKRLKVIDLSVNKIS
PSGDSSEVGFCSNARTSVESYEPQVLEQLHYFRYDKYARSCRFKNKEASFMSVNESCYKYGQ
TLDSLKNSIFFVKSSDFQHLNFLKCLNLSGNLISQTLNGSEFQPLAELRYLDFSNRDLHLH
STAFEEHLHLEVLDISSNSHYFQSEGITHMLNFTKNLKVQLKMMNDNDISSSTSRTESES
LRTLEFRGNHLDVLWREGDNRYLQLFKNLLKLEELDISKNLSFLPSGVFDGMPPNKLKNSL
AKNGLKSFWSKKLQCLKNLETLDLSHNQLTTVPERLSNCSRLKNLILKNNQIRSLTKYFLQ
DAFQLRYLDLSSNKIQMIQKTSFPENVLNNLKMLLLHHNRFLCTCDAVWFVWWVNHTVTTIP
YLATDVTCVGPAGHKGQSVISLDLYTCELDLTNLILFSLISISVSLFMVMMTASHLYFWDVW
YIYHFCKAKIKGYQRLISPDCCYDAFIVYDTKDPVTEWVLAEVLAKLEDPREKHFNLCLEE
RDWLPQGQVLENLSQSIQLSKKTVFVMTDKYAKTENFKIAFYLSHQRLMDEKVDVIIIFLE
KPFQKSFLQLRKRLCGSSVLEWPTNPQAHPPYFWQCLKNALATDNHVAYSQVFKETV

Signal sequence:

amino acids 1-26

Transmembrane domain:

amino acids 840-860

FIGURE 210

GGGTACCATTCTGCGCTGCTGCAAGTTACGGAATGAAAAATTAGAACAACAGAAAACTGGAACACATGTTCTCTTC
AGTCGTCGAATGCTGACCTGCACTTTTCTGCTAAATATCTGGTTCTCTGTGAGTTATGCGCGGAAGAAAAATTTTCTCTA
GAAGCTATCCTGTGTGATGAGAAAAAGCAAAATGACTCAGTTATTGCGAGGTGCAGCAATCGTGCAGTACAGGAAAG
TTCCCAACACGGTGGGCAAAATGATGACAGAACTAGACCTGTCTGATAATTTTCATCACACATAAACGAATGAAT
CATTTCAAGGGTGCAGAAATGCTCATAAAATAAATCTAAACCAACACCCCAATGTACAGCACCAGAACCGGAAATC
CCGGTATACAAATCAAAATGGCTTGAATATCACAGACAGGGGCAATTCCTCAACCTTAAAAAACCCTAAGGGGATTCATGCG
TTGAAGACAACAGTTACCCCAAAATACCCCTCTGGTTTGCCAGAGTCTTTGACAGAACTTAGTCTAAATCAAAAAACA
ATATATACAACTAATCAAAAGAGGGCAATTTCAAGACTTATAAACTTGAAAAATCTCTATTGGCGCTGGAACTGCT
ATTTTAAACAAGTTTGGCAGAAAAATTAACATAGAAGATGGAGTATTGAAACGCTGACAAATTTGGAGTTTGCTAT
CACTATCTTTCAATTTCTTTTCAACGTCGCCACCAACCTGCCAAGCTCCCTACGCAAACTTTTCTGAGCAACA
CCCAGATCAAAATACATTAGTGAAGAGATTTCAAGGGATGATAAATTTAAACATTACTAGATTTAAGCGGGAATC
TCCGAGGTGCTTCAATGCCCATTTCCATGCGTGCCTTTGTGATGGTGGTGCTTCAATTAATATAGATCGTTTGTG
CTTTTCAAAACTTGACCCAACCTTCGATACCTTAAACCTCTCTAGCACTTCCCTCAGGAAGATTAAATGCTGCTGGT
TTAAAAATATGCCTCATCTGAAGGTGCTGGATCTTGAATTCAACTATTTAGTGGGAGAAATAGTCTCTGGGGCAT
TTTAAACGATGCTGCCCCGCTTAGAAAACTTGTACTTGTCTTTTAACTATATAAAGGGAGATTATCCACAGCATATA
TCAGAGAAGATGATTTCCAGCCCTGATGCGAGCTTCCAACTTATCGACTATCAACTGGGTATTAATTTTATTA
AGCAAAATCGATTCTTCAAACTTTCCAAAAATTTCTCAATCTGGAATTTATTTACTTTGTCAAAAAACAATAATCAC
CGTTGGTAAAGATACCCCGCAGAGATTATGCAATAGTTCTCTTTTCAAGCTCATATCCGGAACACAGCCCTCAA
CAGATTTGAGTTTGACCCCAATTCGAACTTTATCAATTCACCCGCTCTTAATAGCTGCAACATGTGCTGCT
ATGGAAGAAGCTTAGATTTAAGCCTCAACAGTATTTCTTCTTGGGCGCAACCAATTTGAAAAATCTTCTGACA
TTGCTGCTTTAAATCTGTCTGCAAAATGCTCAAGTGTAAAGTGAACCTGAATTTTCAAGCTTCTGCTGACT
TCAAAATTTTGGATTTGACAAACAATAGACTAGACTTTTGAATATGCTAGTGTCTTACTGAATTTGCTGCAAT
AAGTTTCTAGATCTCAGTCTAATAATTCACACTATTTAGAAATAGCAGGCGTAAACCATCTAGATTTTCTGAAA
ATTTCAACAATCTAAAGTTTTTAACTTGAGCCACAACAACATTTAATCTTAAACAGATTAAGTATAACCTGGA
GCAAGTCCCTGTAGATAATGATTTTCAAGTGGCAATCGCCTTGAATTTTGTGGAATGATGATGACACAGGATATA
TCTCCATTTTCAAGGTTCTCAAGAACTGACACGCTCGGATTTATCCCTTAAAGTGTGAAGCATCCCAATG
AAGCATCTCTTAAATTTGCCAGCAGTCTCACTGAACTACATAAATGATAATATGTTAAAGTTTTTAACTGGA
CATTACTCCAGCAGTTTCTGCTGCTCGAGTTGCTTGACTTACGTGGAACCAAACTACTCTTTTAACTGATAGCC
TTTCTGAAGTCAGTAGTCTGAAGCACTCGATTTTAAAGTTCCAATCTGCTAAAAACAATCAACAATCCGCACTTG
AAACTAAGACCAACCAAAATTTATCTATGTTGGAACACACGGAACCCCTTTGAATGCACCTGTGACATTTGGAG
ATTTCCGAAGATGAGTGGATGAACATCTGAAATGTCAAAATTTCCAGACTGCTAGATGTCATTTGTGCGAGTCTG
GGATCAAGAGGGAGAGATTTTGTAGTCTGAGGCTAAACAATCTGTGTTTCAAGATGTCACCTGCAGTGATATAT
TTTCTTCAAGTCTTTTATCACCACCATGTTATGTTGGCTGCCCTGGCTCAGATTTGTTTGTGTTGGGATGTT
GGTTTATATAAATGTTGTTTGTAGTAAAGTAAAGGCTACAGGTCTCTTTCCACATCCCAAACTCTTCTATGATG
CTTACATTTCTTATGACACCAAAAGATGCCCTCTGTTACTGACTGGGTGATAAATGAGTGGCGCTACCACTTGAAG
AGAGCCGAGACAAAAACGTTCTCTCTTGTCTAGAGGAGAGGGATTGGGACCCGGGATTTGGCCATCATCGACAAC
TCATGCGAGAGCATCAACCAAGCAAGAAAAACAGTATTTGTTTAAACCAAAAAATATGCAAAAAAGCTGGAACCTTA
AAACAGCTTTTACTTGGCTTTGACAGGCTAATGGATGAGAACATGGATGTGATATATTTTATCTGCTGGAGC
CAGTTGTATCAGCATTTCTCAGTATTTGAGGCTACGCGACGCGATCTGTAAGAGCTCCATCTTCCAGTGGCCCTGACA
ACCCGAAGTCAGAGGCTTTGTTTGGCAACACTGAGAAATGTGTTCTGACTGAAAAATGATCAGGCTATAACA
ATATGTATGTGCAATCCATTAAGCAATTAACCTGACGTTTAAGTCATGATTTGCGGCCATAAATAAGAGTGCAAG
GAATGACATTTCTGTATAGTTATCTATTGCTATGTAACAAATTTATCCCAAACTTAGTGGTTTAAAAACAACACA
TTTGTGCGCCACAGTTTTTGAAGGTCAGGAGTCCAGGCCACGACATAACTGGGTCCTCTGCTCAGGCTGTCTCAG
AGGCTGCAATGTAGGTTTCAACAGAGACATAGGCATCACTGGGGTCACTCATGTGGTTGTTTCTGGATTTCA
ATTTCTCTGGGCTATTTGGCCAGGCTATACCTCATGTAAGCATGCGAGCTCTCCCAACAGGAGGCTTCTGCT
ATCAGAGTCAAAAAAGAGAGGTGCTAGCAAGATGAAGTCAACAATCTTTTGAATATGAAATCAAAAAAGTAAAT
ATCTCATCACTTTGGCCATATTTCTATTTTGAAGTAAACCAAGGTCACCAAGCTTCCAGGAGTCAAC
TCAGTCCAGGGAACAGCTGAGAACCAAGATGTGAGCTGTGATGCTTCAAGTTGGTTCATCAACTATTTTCCCT
TGACTGTGCTCTGGATGGCTGCTATCTTGATGATAGATTTGAAATATGAGGAGGAGGATCACTGTGAGCC
ATCTTAGCAGTTGACCTAACACATCTCTTTTCAATATCTAAGAACTTTTGGCCATCTGACTAATGGCTCAAT
TTAGCTGTGTTTATATTTATCATATATCTAGGCTACATGGTTATATTTGCTGTGGTTCGTTTCTGTTTAT
TTACAGTTGCTTTTCAAAATATTTGCTGTAACTTTGACTTCTAAGTTTGAAGTCCATTTAAGACTGAGATGG
ATAGCTTTTAAAGCATCTTTTACTCTTCAACATTTTAAAAATATGCACTAAATTCGAAGCTTTTGGCTCATATA
TTGTTAATTGCCATTGCTGTAAATCTTAAATGAATGAATAAATGTTTCAATTTTCAAAAAAATAAAAAA

FIGURE 211

MENMFLQSSMLTCIFLLISGSCELCAEENFSRSYPCDEKKQNDSVIAECSNRRLQEVPPQTVG
KYVTELDLSDNFITHITNESFQGLQNLTKINLNHNPNVQHONGNPGIQSNGLNITDGAFLNL
KNLRELLLEDNQLPQIPSGLPESLTELSLIQNNIYNITKEGISRLINLKNLYLAWNCYFNKV
CEKTNIEDGVFETLTNLELLSLSFNSLSHVPPKLPSSLRKLFLSNTQIKYISEEDFKGLINL
TLLDLSGNCPRCFNAPFPVPCDGGASINIDRFQNLTLQLRYLNLSSLSLRKINAAWFKNM
PHLKVLDELFEYNLVGEIVSGAFLTMLPRLEILDLSFNYIKGSPQHINISRNFSKLLSLRAL
HLRGYVQBELREDDFQPLMQLPNLSTINLGINFIKQIDFKLFQNFNSNLEIIYLSENRISPLV
KDTRQSYANSSSFQRHIRKRRSTDFEFDPHSNFYHFTRPLIKPQCAAYGKALDLSLNSIFFI
GPNQFENLPDIACLNLSANSNAQVLSGTEFSAIPHVKYLDLTNNRLDFDNASALTELSDLEV
LDLSYNSHYFRIAGVTHHLEFIQNFTNLKVLNLSHNNIYTLTDKYNLESKSLVELVFSGNRL
DILWNDDDNRYISIFKGLKNLTRLDLSLNRLKHIPNEAFLNLPASLTELHINDNMLKFFNWT
LLQQFPRLELLDLRGNKLLFLTDSLSDFTSSLRTLLLSHNRI SHLPSGFLSEVSSLKHLDL
SNLLKTINKSALETKTTKLSMLELHGNPFECTCDIGDFRRWMDEHLNVKIPRLVDVICASP
GDQRGKSIVSLELTTCVSDVTAIVLFFFTFFITTMVMLAALAHHLFYWDVWFIYNVCLAKVK
GYRSLSTSQTIFYDAYISYDTKASVTDWVINELRYHLEESRDKNVLLCLEERDWDPLGAIID
NLMQSINQSKKTVPVLTKKYAKSWNFKTAFYLALQRLMDENMDVIIIFILLEPVLQHSQYLRL
RQRICKSSILQWPDNPKAEGFLWQTLRNVVLTENDSRYNMYVDSIKQY

Signal sequence:

amino acids 1-26

Transmembrane domain:

amino acids 826-848

FIGURE 212

CCAGGTCCAAC TGACCTCGGTTCTATCGATTGAATTC CCGGGGATCCTCTAGAGATCCCT
CGACCTCGACCCACGCGTCCGCCAAGCTGGCCCTGCACGGCTGCAAGGGAGGCTCCTGTGGA
CAGGCCAGGCAGGTGGGCCTCAGGAGGTGCCTCCAGGCGGCCAGTGGGCCTGAGGCCCCAGC
AAGGGCTAGGGTCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGGCCACGCTGGGC
TCCAGCAGCATCAGCAGCCCCAGGACCGGGGAGGCACAGGTGGCCCCACCA CCGGAGGA
GCAGCTCCTGCCCCCTGTCCGGGGGATGACTGATTCTCCTCCGCCAGGCCACCCAGAGAGAA
GGCCACCCCGCTGGAGGCACAGGCCATGAGGGGCTCTCAGGAGGTGCTGCTGATGTGGCTT
CTGTTGTTGGCAGTGGGCGGCACAGAGCACGCCCTACCGGCCCGGCCGTAGGGTGTGTGCTGT
CCGGGCTCACGGGGACCTGTCTCCGAGTCGTTCTGTGCAGCGTGTGTAC CAGCCCTTCTCTCA
CCACCTGCGACGGGCACCGGCCTGCAGCACCTACCGAACCATCTATAGGACCGCCTACCGC
CGCAGCCCTGGGCTGGCCCTGCCAGGCCTCGCTACGCGTGTGTGCCCCGGCTGGAAGAGGAC
CAGCGGGCTTCTTGGGGCTGTGGAGCAGCAATATGCCAGCCGCCATGCCGGAACGGAGGGA
GCTGTGTCCAGCCTGGCCGTGCCGTGCCCTGCAGGATGGCGGGGTGACACTTGC CAGTCA
GATGTGGATGAATGCAGTGTCTAGGAGGGGCGGCTGTCCCAGCGCTGCATCAACACCGCCGG
CAGTTACTGGTGCCAGTGTGTGGGAGGGGCACAGCCTGTCTGCAGACGGTACACTCTGTGTGC
CCAAGGGAGGGCCCCCAGGGTGGCCCCCAACCGCACAGGAGTGGACAGTGCATGAAGGAA
GAAGTGCAGAGGCTGCAGTCCAGGGTGGACCTGCTGGAGGAGAAGCTGCAGCTGGTGCTGGC
CCCAGTGCACAGCCTGGCCTCGCAGGCATGGAGCATGGGCTCCCGGACCCCGGCAGCCTCC
TGGTGCACCTCCTTCCAGCAGCTCGGCCGCATCGACTCCTTGAGCGAGCAGATTTCTTCTTG
GAGGAGCAGCTGGGGTCTCTGCTCTGCAAGAAAGACTCGTGA CTGCCAGCGCCCCAGGCTG
GACTGAGCCCTCAGCCGCCCTGCAGCCCCCATGCCCTGCCCCAACATGCTGGGGTCCAG
AAGCCACCTCGGGGTGACTGAGCGGAAGGCCAGGCAGGCCCTTCTCTCTTCTCTCTCCCC
TTCTCTCGGAGGCTCCCCAGACCTGGCATGGGATGGGCTGGGATCTTCTCTGTGAATCCAC
CCCTGGCTACCCCCACCTGGCTACCCCAACGGCATCCCAAGGCCAGGTGGGCCCTCAGCTG
AGGGAAGGTACGAGCTCCTGTCTGGAGCCTGGGACCATGGCACAGGCCAGGCAGCCCCGAG
GCTGGGTGGGCCTCAGTGGGGCTGCTGCCTGACCCCCAGCACATAAAAAATGAAACGTGA
AAAGGGCGGCCGCACTCTAGAGT
CGACCTGCAGAAGCTTGGCCGCATGGCCCAACTTGTTTATTGCAGCTTATAATGTTACAAAT

FIGURE 213

MRGSQEVLLMWLLVLAVGGTEHAYRPGRRVCAVRAHGDPVSESFVQRVYQPFLTTCDGHRAC
STYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGACGAATCQPPCRNGGSCVQPGRCR
CPAGWRGDTQCSDVDECSARRGGCPQRCINTAGSYWCQWEGHSLSDGTL CVPKGGPPRVA
PNPTGVDSAMKEEVQRLQSRVDLLEEKQLVLAPLHSLASQALEHGLPDPGSLLVHSFQQLG
RIDSLSEQISFLEEQLGSCSCKKDS

Signal sequence:

1-19

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
21

FIGURE 214

GCCAGGCAGGTGGGCCTCAGGAGGTGCCTCCAGGCGGCCAGTGGGCCTGAGGCCCCAGCAAAG
GGTAGGGTCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGGCCACGCCTGGGCTCC
AGCAGCATCAGAGCAGCCCCCTGTGGTTGGCAGCAAAGTTCAGCTTGGCTGGGCCCGCTGTGA
GGGGCTTCGCGCTACGCCCTGCGGTGTCCCGAGGGCTGAGGTCTCCTCATCTTCTCCCTAGC
AGTGGATGAGCAACCCAAACGGGGGCCCGGGGAGGGGAACTGGCCCCGAGGGAGAGGAACCCC
AAAGCCACATCTGTAGCCAGGATGAGCAGTGTGAATCCAGGCAGCCCCCAGGACCGGGGAGG
CACAGGTGGCCCCCACCACCCGGAGGAGCAGCTCCTGCCCTGTCCGGGGATGACTGATTTC
TCCTCCGCCAGGCCACCCAGAGGAGAAGGCCACCCCGCTGGAGGCACAGGCCATGAGGGGC
TCTCAGGAGGTGCTGCTGATGTGGCTTCTGGTGTGGCAGTGGGCGGCACAGAGCACGCCTA
CCGCCCCGGCCGTAGGGTGTGTGCTGTCCGGGCTCACGGGGACCTGTCTCCGAGTCGTTTCG
TGACGCTGTGTACAGCCCTTCTCTACCACCTGCGACGGGCACCGGCTCGCAGCACCTAC
CGAACCATCTATAGGACCGCTACCGCCGACAGCCCTGGGCTGGCCCCCTGCCAGGCCTCGTA
CGCGTGTGCCCCGCTGGAAGAGGACCAGCGGGCTTCTGGGGCTGTGGAGCAGCAATAT
GCCAGCCGCCATGCCGGAACGGAGGGAGCTGTGTCCAGCCTGGCCGTGCCGTGCCCTGCA
GGATGGCGGGGTGACACTTGCCAGTCAGATGTGGATGAATGCAGTGTAGGAGGGGCGGCTG
TCCCCAGCGCTGCATCAACACCGCCGGCAGTTACTGGTGCCAGTGTGTGGAGGGGCACAGCC
TGTCTGCAGACGGTACACTCTGTGTGCCCAAGGGAGGGCCCCCAGGGTGGCCCCCAACCCG
ACAGGAGTGGACAGTGAATGAAGGAAGAAGTGCAGAGGCTGCAGTCCAGGGTGGACCTGCT
GGAGGAGAAGCTGCAGCTGGTGTCTGGCCCCACTGCACAGCCTGGCCTCGCAGGCACCTGGAGC
ATGGGCTCCCGGACCCCGGCAGCCTCTGGTGCACTCCTTCCAGCAGCTCGGCCGCATCGAC
TCCCTGAGCGAGCAGATTTCTTCTGGAGGAGCAGCTGGGGTCTGCTCCTGCAAGAAAGA
CTCGTGACTGCCAGCGCTCCAGGCTGGACTGAGCCCCCTACGCCGCCCTGCAGCCCCCATG
CCCCTGCCCAACATGCTGGGGGTCCAGAAGCCACCTCGGGGTGACTGAGCGGAAGGCCAGGC
AGGGCCTTCTCCTCTTCTCCTCCCTTCTCGGGAGGCTCCCCAGACCTGGCATGGGAT
GGGCTGGGATCTTCTGTGAATCCACCCCTGGCTACCCCCACCTGGCTACCCCAACGGCA
TCCCAGGCCAGGTGGACCCTCAGCTGAGGGAAGGTACGAGCTCCTGCTGGAGCCTGGGAC
CCATGGCACAGGCCAGGCAGCCCGAGGCTGGGTGGGGCCTCAGTGGGGCTGCTGCCTGAC
CCCCAGCACAAATAAAATGAAACGTG

FIGURE 215

MRGSQEVLLMWLLVLAVGGTEHAYRPGRRVCAVRAHGDPVSESFVQRVYQPFLTTCDGHRAC
STYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGACGAATCQPPCRNGGSCVQPGRCR
CPAGWRGDTQCSDVDECSARRGGCPQRCINTAGSYWCQWEGHSLSDGTL CVPKGGPPRVA
PNPTGVD SAMKEEVQRLQSRVDLLEEKQLVLAPLHSLASQALEHGLPDPGSLLVHSFQQLG
RIDSLSEQISFLEEQLGSCSCKDS

Signal sequence:

1-19

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
219

FIGURE 216

CCCACGCGTCCGAAGCTGGCCCTGCACGGCTGCAAGGGAGGCTCCTGTGGACAGGCCAGGCA
GGTGGGCTCAGGAGGTGCCTCCAGGCGGCCAGTGGGCTGAGGCCCCAGCAAGGGCTAGGG
TCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGGCCACGCCCTGGGCTCCAGCAGCAT
CAGCAGCCCCCAGGACCGGGGAGGCAAGGTGGCCCCACACCACCCGAGGAGCAGCTCCTGC
CCTGTCCGGGGGATGACTGATTCTCCTCCGCCAGGCCACCCAGAGGAGAAGGCCACCCCGC
CTGGAGGCACAGGCCATGAGGGGCTCTCAGGAGGTGCTGCTGATGTGGCTTCTGGTGTGGC
AGTGGGCGGCACAGAGCACGCCTACCGCCCGGCCGTAGGGTGTGTGCTGTCCGGGCTCAGC
GGGACCTGTCTCCGAGTCGTTTCGTGCAGCGTGTGTACCAGCCCTTCTCACCACCTGCGAC
GGGCACCGGGCTGCAGCACCTACCGAACCATCTATAGGACCGCCTACCGCCGAGCCCTGG
GCTGGCCCCCTGCCAGGCCTCGCTACGCGTGCTGCCCCGGCTGGAAGAGGACCAGCGGGCTTC
CTGGGCTGTGGAGCAGCAATATGCCAGCCGCCATGCCGGAACGGAGGGAGCTGTGTCCAG
CCTGGGCGCTGCCGCTGCCCTGCAGGATGGCGGGTGACACTTGCCAGTCAGATGTGGATGA
ATGCAGTGCTAGGAGGGGCGGCTGTCCCCAGCGTGCGTCAACACCGCCGGCAGTTACTGGT
GCCAGTGTGGGAGGGGCACAGCCTGTCTGCAGACGGTACACTCTGTGTGCCCAAGGGAGGG
CCCCCAGGGTGGCCCCCAACCCGACAGGAGTGGACAGTGCAATGAAGGAAGAAGTGACAG
GCTGCAGTCCAGGGTGGACCTGCTGGAGGAGAAGCTGCAGCTGGTGTGCCCCACTGCACA
GCCTGGCCTCGCAGGCACTGGAGCATGGGCTCCCGGACCCCGGCAGCCTCCTGGTGCACCTC
TTCAGCAGCTCGGCCGCATCGACTCCCTGAGCGAGCAGATTCTCTTCTGGAGGAGCAGCT
GGGGTCTGCTCCTGCAAGAAAGACTCGTGACTGCCAGCGCCCCAGGCTGGACTGAGCCCC
TCACGCCGCCCTGCAGCCCCCATGCCCTGCCCAACATGCTGGGGTCCAGAAGCCACCTCG
GGGTGACTGAGCGGAAGGCCAGGCAGGGCCTTCTCCTCTTCTCCTCCTCCCTTCTCGGGAG
GCTCCCCAGACCCTGGCATGGGATGGGCTGGGATCTTCTCTGTGAATCCACCCTGGCTACC
CCCACCCTGGCTACCCCAACGGCATCCCAAGGCCAGGTGGGCCCTCAGCTGAGGGAAGGTAC
GAGTCTCCCTGCTGGAGCCTGGGACCCATGGCACAGGCCAGGCAGCCGGAGGCTGGGTGGG
CCTCAGTGGGGCTGCTGCCTGACCCCCAGCACATAAAAATGAAACGTG

FIGURE 217

MRGSQEVLLMWLLVLAVGGTEHAYRPGRRVCAVRAHGDPVSESFVQRVYQPFLTTCDGHRAC
STYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGACGAATCQPPCRNGGSCVQPGRCR
CPAGWRGDTQCSDVDECSARRGGCPQRCVNTAGSYWCQWEGHSLSDGTL CVPKGGPPRVA
PNPTGVD SAMKEEVQRLQSRVDLLEEKQLVLAPLHSLASQALEHGLPDGSLLVHSFQQLG
RIDSLSEQISFLEEQLGSCSCKKDS

Signal sequence:

1-19

1000
900
800
700
600
500
400
300
200
100
0
100
200
300
400
500
600
700
800
900
1000

FIGURE 218

GGTTGCCACAGCTGGTTTAGGGCCCCGACCACTGGGGCCCTTGTTCAGGAGGAGACAGCCTCCGGGCCCGGGGAG
GACAAGTCGCTGCCACCTTTGGCTGCCGACGTGATTCCCTGGGACGCTCGTTTCTCGCGTCAGCTGCCGGCCG
AGTTGGGTTCCCGTGGTTTCAGGCCGGCTCCCCCTTCCTGTGCTCCCTTCTCCCGCTGGGGCCGGTTTATCGGGAGG
AGATTGTCTTCACGGCTAGCAATTTGACCTTTTGATGATGTTTGACCCAGCGCAGGAATAGCAGGCAACGTGAT
TTCAAAGCTGGGCTCAGCCTCTGTTTCTTCTCTCGTGAATCGCAAAACCCATTTTGGAGCAGGAATTCACATCA
TGTCTGTGATGGTGGTGAGAAAGAGGTGACACGGAAATGGGAGAACTCCAGGCAGGAACACCTTTTGCTGTG
ATGGCCGCGTCATGATGGCCCGGCAAAAGGGCATTTTCTACCTGACCCTTTTCTCATCTCGGGGACATGTACAC
TCTTCTTGCGCTTTGAGTGCCGCTACCTGGCTGTTGAGCTGTCTCTGCCATCCCTGTATTGTCTGCCATGCTCT
TCCCTTTCTCATGGCTACACTGTTGAGGACCAGCTTCAGTGACCCTGGAGTGATTCTCGGGCGCTACCGATG
AAGCAGCTTTCTAGAAATGGAGATAGAAAGCTACCAATGGTGGCTGCCCGGGCCAGCGACACCGCCCTCGTA
TCAAGAATTTCCAGATAAAACACAGATTGTGAAACTGAAATACCTGTACACATGCAAGATCTTCCGGCCCTCCCC
GGGCTCCCATTTGCAGCATCTGTGACAACTGTGTGGAGCGCTTCGACCATCACTGCCCTCGGGTGGGGAATTGTG
TTGAAAAGAGGAACTACCGCTACTTCTACCTCTTCATCCCTTCTCTCTCCCTCTCAACAATCTATGCTCTCGCT
TCAACATCGTCTATGTGGCCCTCAAATCTTTGAAAATTGGCTTCTTGAGACATTGAAAGAACTCTGGAACCTG
TTCTAGAAGTCTCATTTGCTTTTACACTCTGGTCCGTCTGTGGGACTGACTGGATTTCATCTTTCTCGTGG
CTCTCAACAGACAAACCAATGAAGACATCAAGGATCATGGACAGGGAAGAAATCGCGTCCAGAATCCCTACAGCC
ATGGCAATATTGTGAAGAACTGCTGTGAAGTGCTGTGTGGCCCTTGGCCCGCAGTGTGCTGGATCGAAGGGGTA
TTTTGCCACTGGAGGAAAGTGAAGTCGACCTCCAGTACTCAAGAGACAGTAGCAGCCTCTTGCCAACAGACC
CAGCCCCCAGAAACCTGAACTCAAAATGAGATGCCGGAGGACAGCAGCACTCCGAGAGATGCCACCTCCAG
AGCCCCCAGAGCCACCCAGGAGGCGAGCTGAAGCTGAGAAGTAGCCTATCTATGAAGAGACTTTTGTGTGTGT
TAATTAGGGCTATGAGAGATTTCAGTGAGAAGTTAAACCTGAGACAGACAGCAAGTAACTGTCTCCCTTTTAACT
GTTTTTCTTTGGCTTTTATTCACCCAGTTGACACACTGCACTTTCTGTCTGCAAGCTTTTAAATTTCTGAACT
CAAGGCGTGGCAGAGATGTCAGTCACTCTGATAACTGGAATAATGGGCTCTTTGGGCCCTGGCACTGGTTCT
CGATGGCTCAGCCACAGGTCCTTGGACCCCTCTCTTCCCTCAGATCCAGGCCCTCTGTCTGGGGTCACT
TGGTCTCATTTCTGGGCTAAAGTTTGTGAGACTGGCTCAAACTCTCCCAAGCTGTGACGTGCTGAGTCCAGA
GGCAGTCAAGAGACCTCTGGCCAGGAGTCCAACTGGGTTCTTTGGGGTCTTCAGGACTGAAGAGGAGGGAGAG
TGGGCTCAGAAGATTCTCTGGCCACCAAGTGCCAGCATTGCCCAAACTCTTTAGGAATGGGACAGGTACTT
TCCACTGTGTGTANNNNNNNNNNNNNNNNNNNNNNNNNNNNTTGTCTTCTTTGACTCTGCTCCCATTTAGGAG
CAGGAATGGCAGTAAATAAAGTCTGCACTTTGGTCACTTTCTTTTCCAGAGGAAGCCGAGTGCTCACTTAAAC
ACTATCCCTCAGACTCCCTGTGTGAGGCTGACAGGGCCCTGAATGCACAAATGGGAAACCAAGGCACAGAGAG
GCTCTCTCTCTCTCTCTCTCCCCGATGATCCCTCAAAAAAATAAATGCTAAACAGATTCTTCCATTAAAGCT
CGGCTGAGTGAGGGAAGCCAGCACTGCTGCCCTCTCGGGTAACTACCCCTAAGGCCTCGGCCCACTCTGGCT
ATGGTAACCACTCGGGGCTTCTCCAGCCCGCTCTTCCAGCACTTCCACCGGCAGAGTCCAGAGCCACTT
CACCTCGGGGTGGCTGTGSCCCCGATGAGCTGTCTCAGGACCTGTCTATTTCAGGGAAGAAGATTATGT
ATTATATGTGGCTATTTCCTAGAGCACTGTGTTTCTCTTTCTAAGCCAGGCTCGGTCTGATGACTTAT
CGGTGGGGGAGTGTAAACCGGAACCTTTCTATCTATTGAAGCGGATTAACCTGTGTCAATGCA

FIGURE 219

MSVMVVRKKVTRKWEKLPGRNTFCDDGRVMMARQKGI FYLTFLILGTCTLFFAFECRYLAV
QLSPAIPVFAAMLFLFSMATLLRTSFSDPGVIPRALPDEAAFIEMEIEATNGAVPQGQRPPP
RIKNFQINNQIVKLKYCYTCKIFRPPRASHCSICDNCVERFDHHC PWVGN CVGKRNYRYFYL
FILSLSLTIYVFAFNIVYVALKSLKIGFLET LKETPGTVLEVLCFFTLWSVVGLTGFHTF
LVALNQT TNEDIKGSWTGKNRVQNPYSHGNIVKNCCVLCGPLPPSVLDRRGILPLEESGSR
PPSTQETSSSLLPQSPAPTEHLNSNEMPEDSSTPEEMPPEPPEPPQEAAEA EK

Putative transmembrane domains:

amino acids 36-55 (type II TM), 65-84, 188-208, 229-245

FIGURE 220

AAAACCCCTGTATTTTTTACAATGCAAATAGACAATNANCCTGGAGGTCTTTGAATTAGGTAT
TATAGGGATGGTGGGGTTGATTTTTTNTTCCTGGAGGCTTTTGGCTTTGGACTCTCNCTTTCT
CCCACAGAGCNCTTCGACCATCACTGCCCTGGGTGGGGAATTGTGTTGGAAGAGGAACTA
CCGCTANTTCTACCTCTTCATCCTTTNTCTCTCCNCCTCACAATCTATGCTTCGCCTTCA
ACATCGT

FIGURE 221

GTTGTGTCCTTCAGCAAAACAGTGGATTAAATCTCCTTGCAAGCTTGAGAGCAACACAA
TCTATCAGGAAAGAAAGAAAAAACCGAACCTGACAAAAAGAAGAAAAAGAAGA
AAAAAATCATGAAAACCATCCAGCCAAAAATGCACAATTCTATCTCTTGGGCAATCTTCA
GGGCTGGCTGCTCTGTGTCTCTTCCAAGGAGTGCCCGTGCGCAGCGGAGATGCCACCTTCC
CCAAAGCTATGGACAACGTGACGGTCCGGCAGGGGAGAGCGCCACCCTCAGGTGCATATT
GACAACCGGTCACCCGGGTGGCTGGCTAAACCGCAGCACCATCTCTATGCTGGGAATGA
CAAGTGGTGCTGGATCCTCGCGTGGTCTTCTGAGCAACACCCAAACGCAGTACAGCATCG
AGATCCAGAACGTGGATGTGTATGACGAGGGCCCTTACACCTGCTCGGTGCAGACAGACAAC
CACCCAAAGACCTCTAGGGTCCACCTCATTTGTGCAAGTATCTCCCAAATTTAGAGATTTC
TTCAGATATCTCCATTAATGAAGGGAACAATATTAGCCTCACCTGCATAGCAACTGGTAGAC
CAGAGCCTACGGTTACTTTGGAGACACATCTCTCCAAAGCGTTGGCTTTGTGAGTGAAGAC
GAATACTTTGGAATTCAGGGCATACCCGGGAGCAGTCAGGGGACTACGAGTGCAGTGCCCTC
CAATGACGTGGCCGCGCCCGTGGTACGAGAGTAAGGTACCGTGAACTATCCACCATACA
TTTCAGAAGCCAAGGTTACAGGTGTCCCGTGGGACAAAAGGGGACACTGCAGTGTGAAGCC
TCAGCAGTCCCTCAGCAGAATTCAGTGGTACAAGGATGACAAAAGACTGATTGAAGGAAA
GAAAGGGGTGAAAGTGGAAAACAGACCTTTCTCTCAAAACTCATCTTCTTCAATGTCTCTG
AACATGACTATGGGAACTACACTTGCGTGGCTCCAACAGCTGGGCCACACCAATGCCAGC
ATCATGCTATTTGGTCCAGGCGCCGTGAGCAGGTGAGCAACGGCAGTCGAGGAGGGCAGG
CTGCGTCTGGCTGCTGCCCTCTTCTGGTCTTGACCTGCTTCTCAAATTTTGAATGTGAGTGCC
ACTTCCCCACCCGGGAAAGGCTGCCGCCACCACCACCAACACAACAGCAATGGCAACAC
CGACAGCAACCAATCAGATATATACAAATGAAATTAGAAGAAACACAGCCTCATGGGACAGA
AATTTGAGGGAGGGGAACAAAGAATACTTTGGGGGAAAAGAGTTTTAAAAAAGAAATTGAA
AATTGCCTTGACAGATATTTAGGTACAATGGAGTTTCTTTTCCCAAACGGGAAGAACACAGC
ACACCCGGCTTGGACCCACTGCAAGCTGCATCGTGCAACCTCTTTGGTGCCAGTGTGGGCAA
GGGCTCAGCCTCTCTGCCCCAGAGTGCCCCACGTGGAACATTCTGGAGCTGGCCATCCCA
AATTCAATCAGTCCATAGAGACGAACAGAATGAGACCTTCCGGCCCAAGCGTGGCGCTGCGG
GCACTTTGGTAGACTGTGCCACCACGGCGTGTGTTGTGAAACGTGAAATAAAAAGAGCAAAA
AAAAA

Signal peptide:
amino acids 1-28

Signal peptide:

amino acids 1-28

GAAAAAAATCATGAAAAACCATCCAGCCAAAAATGCACAATTCTATCTCTTGGGCAATCTTC
ACGGGGCTGGCTGCTCTGTGTCTCTTCCAAGGAGTGCCCGTGCGCAGCGGAGATGCCACCTT
CCCCAAGCTATGGACAACGTGACGGTCCGGCAGGGGGAGAGCGCCACCCTCAGGTGACTA
TTGACAACCGGGTCACCCGGGTGGCTTGGCTAAACCGCAGCACCATCCTCTATGCTGGGAAT
GACAAGTGGTGCTTGGATCCTCGCGTGGTCTTCTGAGCAACACCCAAACGCAGTACAGCAT
CGAGATCCAGAACGTGGATGTGTATGACGAGGGCCCTTACACCTGCTCGGTGCAGACAGACA
ACCACCCAAAGACCTCTAGGGTCCACCTCATTTGTGCAAGTATCTCCCAAAATTGTAGAGATT
TCTTCAGATATCTCCATTATGAAGGGAACAATATTAGCCTCACCTGCATAGCAACTGGTAG
ACCAGAG

FIGURE 224

ATGGCTGGTGACGGCGGGGCGGGGACGGGGGCGGGGCGGGGAGCGGGCAGCTGCGGGAGCCCTGA
 ATCACCGCTGGCCGAGCTTCACACATGAACCTGCGCGCTGCAGGAGCTGGGAGCTGGCAGCAAGTGGGATTCCAG
 AAGGGACAAGACAGAGCTGTAGGCTCACGACGAGCTGGAGCTGGTCTTAGCAGGTGCTCTCTACTGCTGGCT
 GCACTGCTTCTGGGCTGCTTGTGGCCCTAGGGGTCCAGTACCAACAGAGACCCATCCACAGACCTGCTTACA
 GAGGCTGCATTCCAGTGGCTGGAAAAATCTTGGAGTCCCTGGACCGAGGGGTGAGCCCTGTGAGGACTTTTAC
 CAGTTCTCTTGGGGGTGAGTTCCGAGGAACCCCTGCGGATGGGCGTTCTGCTGGAACACCTTCAACAGC
 CTCTGGGACCAAAACAGGCCATATGAAGCACCTGCTTGAACACCACTTCAACTCCAGCAGTGAAGCTGAG
 CAGAAGACACAGCGCTTCACTTATCTTGCCTACAGGTGGAGCGCATTGAGGAGCTGGGAGCCAGCCAATGAGA
 GACCTCATTGAGAAGATTGGTGGTTGGAACATTACGGGGCCCTGGGACAGGACCACTTTATGGAGGTGTGAAG
 GCAGTAGCAGGGACCTACAGGGCCACCCATTTCTTACCGTCTACATCAGTGCAGCTTAAAGATTCCAACAGC
 AATGTTATCCAGGTGGACAGCTCTGGGCTCTTCTGCGCTCTCGGGATTACTACTTAAACAGAACTGCCAATGAG
 AAAGTGTCTCACTGCTATCTGGATTACATGGAGGAACCTGGGGATGCTGCTGGGTGGGCGGCCACCTCCACGAGG
 GAGCAGATGCAGCAGGTGCTGGATTGGAGATACAGCTGGCCAAATCAGAGTGGCCAGGACAGCGCGCGAC
 GAGGAGAAGATCTACCACAAGATGAGCATTTCGGAGCTGCAGGCTCTGGCGCCCTCCATGGACTGGCTTGAGTTC
 CTGTCTTCTTGTCTGACCAATTGGAGTTGAGTGACTCTGAGCCCTGTGGTGGTGTATGGGATGGATTATTGACAG
 CAGGTGTGACAGCTCATCAACCGCAGGGAACCAAGCATCTGAACAATTACCTGATCTGGAACCTGGTGCAAAAG
 ACAACCTCAAGCCTGGACCGACGCTTTGAGTCTGCACAAGAGAAGCTGCTGGAGACCTCTATGGCACAAGAAG
 TCCTGTGTGCGAGGTGGCAGACCTGCATCTCCAACACGGATGAGCGCCCTGGCTTTGCTTTGGGTCACTCTTTC
 GTGAAGGCCACGTTTGAACCGCAAAAGCAAGAAATTCAGAGGGGATGATCAGCGAAATCCGACACGCAATTGAG
 GAGGCCCTGGGACAGCTGGTTGGATGGATGAGAAGACCCGACGAGCAGCAAGGAGAAGCAGATGCCATCTAT
 GATATGATGGTTTCCAGACTTTATCTTGGAGCCCAAGAGCTGGATGATGTTTATGACGGGTACGAAATTTCT
 GAAGATTCTTTCTTCCAACATGTTGAATTTGTACAACCTCTCTGCAAGGTTTATGGCTGACCACTCCGCAAG
 CCTCCAGCCGAGACAGGTGGAGCATGACCCCCAGACAGTGAATGCCCTACTACTTCCAACCTAAGAATGAGATC
 GTCTTCCCGCTGGCATCTGCGAGGCCCTTCTATGCGCGCAACCAACCCAGGCCCTGAACTCCGTGGCATC
 GGTGTGGTCTATGGGCATGAGTTGACGCAATGCCCTTGATGACCAAGGGCGGAGATGACAAAGAAGGGAACCTG
 CGGCCCTGGTGGCAGATGAGTCCCTGGCAGCTTCCGGAACCAACACGGCTGCAATGGAGGAACAGTCAATCAA
 TACCAGGTCAATGGGAGAGGCTCAACGGCCGCGCAGACGCTGGGGAGAACATTACTGACACAGGGGGCTGAAG
 GCTGCCTACAATGCTTACAAGCATGGCTGAGAAAGCATGGGGAGGAGCAGCAACTGCCAGCGGTGGGGCTCACC
 AACCAACAGCTCTTCTTGTGGATTGCCCCAGTGTGGTGTCTGGTCCGCAACACAGAGAGCTCTCACGAGGG
 CTGGTGAACGACCCCAACAGCCCTGCGCGCTTCCGCGTGTGGGCACTCTTCCAACCTCCGTGACTTCTTGGG
 CACTTGGCTGCGCTGTGGCTCCCCATGAACCCAGGGCAGCTGTGTGAGGTGTGGTGAACCTGGATCAGGGGA
 GAAATGCGCAGCTGTCACCAAGCTGGGGCAGCTCTCTGACAAAGCTGTTGCTCTTGGGTTGGGAGGAAGCAA
 ATGCAAGCTGGCTGGGTCTAGTCCCTCCCCCCACAGGTGACATGAGTACAGACCTCTCAATCAGCACTG
 TGCTCTGCTTTGGGGGTGCCCCGTGCTCCAGCAGAGCCCCACCAATCACTGTGACATCTTCCGTGTACCCCT
 GCGTGAAGAGGTCTGGGTGGGAGGCCAGTTCCCATAGGAAGAGTCTGCG

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984

FIGURE 225

MNVALQELGAGSNVGFQKGTRQLLGSRTQLELVLAGASLLAALLLGCLVALGVQYHRDP SH
STCLTEACIRVAGKILES LDRGVSPCEDFYQFSCGWI RRNPLPDGRSRWNTFNSLWDQQA
ILKHLENTTFNSSSEAEQKTQRFYLSCLQVERIEELGAQPLRDLIEKIGGWNITGPWDQDN
FMEVLKAVAGTYRATPFFT VYISADSKSSNSNVIQVDQSGFLFLPSRDYYLNRTANEKVLTAY
LDYMEELGMLLGGRPTSTREQMQQVLELEIQLANITVPQDQRRDEEKIYHKMSISELQALAP
SMDWLEFLSFLLSPLELSDSEPVVVYGMDYLQQVSELINRTEPSILNNYLIWNLVQKTTSSL
DRRFESAQEKLETTYGTTKSCVPRWQTCISNTDDALGFALGSLFVKATFDRQSKEIAEGMI
SEIRTA FEALGQLVWMDKTRQAAKEKADAIYDMIGFPDFILEPKELDDVYDGYEISEDSF
FQNMNLNLYNFSAKVMADQLRKPPSRDQWSMTPQTVNAYYLP TKNEIVFPAGILQAPFYARNH
PKALNFGGIGVVMGHELTHAFDDQGREYDKEGNLRPWWQNESLAAFRNHTACMEEQYNQYQV
NGERLNGRQTLGENITDNGGLKAAYNAYKAWLRKHGEEQQLPAVGLTNHQLFFVGFAQVWCS
VRTPESSHEGLVTDPHSPARFVRLGTLSNSRDLRHF GCPVGS PMNPGQLCEVW

Type II Transmembrane domain:

amino acids 32-57

FIGURE 227

GGCCGAGCGGGGTGCTGCGCGCGGCCGTGATGGCTGGTGACGGCGGGGCCGGGCAGGGGA
CCGGGGCCGCGCCCGGGAGCGGGCCAGCTGCCGGGAGCCCTGAATCACCGCCTGGCCCGAC
TCCACCATGAACGTGCGCTGCAGGAGCTGGGAGCTGGCAGCAACGTGGGATTCCAGAAGGG
GACAAGACAGCTGTTAGGCTCACGCACGCAGCTGGAGCTGGTCTTAGCAGGTGCCTCTCTAC
TGCTGGCTGCACTGCTTCTGGGCTGCCTTGTGGCCCTAGGGGTCCAGTACCACAGAGACCCA
TCCCACAGCACCTGCCTTACAGAGGCCTGCATTGAGTGGCTGGAAAAATCCTGGAGTCCCT
GGACCGAGGGGTGAGCCCCGTGTGAGGACTTTTACCAGTTCTCCTGTGGGGGCTGGATTGGA
GGAACCCCTGCCCCGATGGGCGTTCTCGCTGGAACACCTTCAACAGCCTCTGGGACAAAAAC
CAGGCCATACTGAAGCACCTGCTTGAAAAACACCACCTTCAACTCCAGCAGTGAAGCTGAGCA
GAAGACACAGCGCTTCTACCTATCTTGCCTACAGGTGGAGCGCATTGAGGAGCTGGGAGCCC
AGCCACTGAGAGACCTCATTGAGAAGATTGGTGGTTGGAACATTACGGGGCCCTGGGACCAG
GACAACTTTATGGAGGTGTTGAAGGCAGTAGCAGGGACCTACAGGGCCACCCCATTTCTTAC
CGTCTACATCAGTGCCGACTCTAAGAGTTCCAACAGCAATGTTATCCAGGTGGACCAGTCTG
GGCTCTTTCTGCCCCCTCGGGATTACTACTTAAACAGAAGTCCCAATGAGAAAGTAAGGAAC
ATCTTCCGAACCCCATCCCTACCCCTGGCTGAGCTGGGCTGATCCCTGTTGACTTTTCCCT
TTGCCAAGGGTCAGAGCAGGGAAGGTGAGCCTATCCTGTCACTAGTGAACAACTGCCCCCT
CCTTCTTTCTTCTTTTCTTCTCCCTCCCTCCCTTTCTTCCCTTTTCTTCTTCTTCTTCC
TCTTATTCTTCTAGTAGGTTTCATAGACCTACTGTGTGCCAGGTCCAGTGGGGGAATTG
GAGATATAAGTTTCCGAGCCATTGCCACAGGAAGCGTTTCAAGTGTGATGGGTTTCATGGACCT
AGATAGGCTGATAACAAAGCTCACAAGAGGGTCTGAGGATTGAGGAGACTTATGAGGCC
AGCAAAGTCTTCTGAAGAGATTGCATTTGAGCCAGGTCCTGTAG

FIGURE 228

ATGCCTACTACCTTCCAAC TAAGAATGAGATCGTCTTCCCCGCTGGCATCCTGCAGGCCCCC
TTCTATGCCCGCAACCACCCCAAGGCCCTGAAC TTGGTGGCATCGGTGTGGTCATGGGCCA
TGAGTTGACGCATGCC TTTGATGACCAAGGGCGCGAGTATGACAAAGAAGGGAACCTGCCGC
CCTGGTGGCAGAATGAGTCCC TGGCAGCCTTCCGGAACCAACGCGCTGCATGGAGGAACAG
TACAATCAATACCAGGTCAATGGGGAGAGGCTCAACGCGCCGCCAGACGTGGGGGAGAACAT
TGCTGACAAACGGGGGCTGAAGGCTGCCTACAATGCTTACAAAGCATGGCTGAGAAAGCATG
GGGAGGAGCAGCAACTGCCAGCCGTGGGGCTCACCAACCACCAGCTCTTCTTCGTGGGATT
GCCCAGGTGTGGTGCTCGGTCCGCACACCAGAGAGCTCTCACGAGGGGCTGGTGACCGACCC
CCACAGCCCTGCCCGCTTCCGCGTGCTGGGCAC TCTCTCCAAC TCCCGTGA CTTCCTGCGGC
ACTTCGGCTGCCCTGTGCGCTCCCCCATGAACCCAGGGCAGCTGTGTGAGGTGTGGTAGACC
TGGATCAGGGGAGAAATGGCCAGCTGT CACCAGACCTGGGGCAGCTCTCCTGACAAAGCTGT
TTGCTCTTGGGTTGGGAGGAAGCAAATGCAAGCTGGGCTGGGTCTAGTCCCTCCCCCCCACA
GGTGACATGAGTACAGACCCCTCCTCAATCACCACATGTGCGCTCTGCTTTGGGGGTGCCCT
GCCTCCAGCAGAGCCCCCACCATTCACTGTGACATCTTTCCGTGTACCCCTGCCTGGAAGAG
GTCTGGGTGGGGAGGCCAGTTC C C C C A T A G G A A G G A G T C T G C C T C T T G T C C C C A G G C T C A C T
CAGCCTGGCGGCCATGGGGCCTGCCGTGCCTGCCCCACTGTGACCCACAGGCCTGGGTGGTG
TACCTCCTGGACTTCTCCCCAGGCTCACTCAGTGCGCACTTAGGGGTGGACTCAGCTCTGTC
TGGCTCACCCTCAGGGCTACCCCCACCTCACCCTGTGCTCCTTGTGCCACTGCTCCCAGTG
CTGCTGCTGACCTTCACTGACAGCTCCTAGTGGAAGCCCAAGGGCTCTGAAAGCCTCCTGC
TGCCCACTGTTTCCCTGGGCTGAGAGGGGAAGTGCATATGTGTAGCGGTACTGGTTCTGT
GTCTTAGGGCACAAGCCTTAGCAAATGATTGATTCTCCCTGGACAAAGCAGGAAAGCAGATA
GAGCAGGGAAAAGGAAGACAGAGTTTATTTTACAGAAAAGAGGGTGGGAGGGTGTGGTCT
TGGCCCTTAGGACC

FIGURE 229

CCCACGCGTCCGAGCCGCCGAGAATTAGACACACTCCGGACGCGGCCAAAAGCAACCGAGA
 GGAGGGGAGGCAAAAACACCGAAAAACAAAAGAGAGAAACAAACCCCAACACTGGGGTGG
 GGGGAGAAAGAAAGAAAGAAACCCACCCACCCACCAAAAAAAGAAAAAAGAAAAA
 AAAAAAATAATCCTGTGGCGCGCGCCTGGTTCCCGGGAAGACTCGCGACGACAGGGGG
 TGGGGGAGTGCAGAGCTGAAAGCTGCTGGAGAGTGAGCAGCCCTAGCAGGGATGGACATGATG
 CTGTTGGTGCGAGGTGCTTGTGTGCGAACCGATGGCTGGCGGCGGTGCTCCTCAGCCTGTG
 CTGCTGTACCCCTCCTGCCTCCCGCTGGACAGAGTGTGGACTTCCCTGGGCGCGGTGG
 ACAACATGATGGTCAGAAAGGGGACACGGCGGTGCTTAGTGTTATTGGAAGATGGAGCT
 TCAAAGGTGCTGGCTGCAACCGGTCAAGTATTATTTTGGCGGAGGTGATAAGTGGTCAGT
 GGATCCTCGAGTTTCAATTTCAACATTGAATAAAGGGACTACAGCCTCCAGATACAGAATG
 TAGATGTGACAGATGATGGCCCATACAGTGTCTGTGTCAGACTCAACATACCCGAGAACA
 ATGCAGGTGCATCTAACTGTGCAAGTTCCTCTAAGATATATGACATCTCAAATGATATGAC
 CGTCAATGAAGGAACCAACGTCACTCTTACTTGTGTTGGCCACTGGGAAACAGAGCCCTCCA
 TTTCTTGGCGACACATCTCCCATCAGCAAAACCATTTGAAAATGGACAATATTTGGACATT
 TATGGAATTACAAGGGGACGCTGGGGAATATGAATGCAGTGCAGAAATGCGTGTCTATT
 CCCAGATGTGAGGAAGAATAAAGTTGTTGTCAACTTGGCTCCTACTATTCAGGAAATTAAT
 CTGGCACCGTGACCCCGGACGCACTGGCCTGATAAGATGTGAAGGTGCAGGTGTGCCGCT
 CCAGCCTTTGAATGGTACAAAGGAGAGAGAAGACTCTTCAATGGCCAAACAGGAATTTATTAT
 TCAAATTTTAGCACAAGATTCATTCTCACTGTTACCAAGCTGACACAGGAGCACTTCGGCA
 ATTATACCTGTGTGGCTGCCAACAGCTAGGCACAACCAATGCGAGCCTGCCTCTTAACCTT
 CCAAGTACAGCCAGTATGGAATTACCGGGAGCGCTGATGTTCTTTTCTCCTGCTGGTACCT
 TGTGTTGACACTGTCTCTTTTACCAGCATATCTACCTGAAGAAATGCCATTCTACAATAAA
 TTTCAAAGACCCATAAAAGGCTTTTAAGGATTCTCTGAAAGTGTCTGAGGTGGATCCAATCT
 GGTACAGTTTGTAAAGCAGCGTGGGATATAATCAGCAGTGTCTTACATGGGGATGATCGCC
 TTTCTGTAGAATTGCTCATTATGTAATACTTTAATTCTACTCTTTTGTATTAGCTACATTA
 CCTTGTGAAGCAGTACACATTGCTCTTTTAAAGCGTGAAGCTCTGAAATTACTTTTAG
 AGGATATTAATTGTGATTTTCATGTTTGTAACTCAACCTTTTCAAAGCACTTCAGTCATGGT
 CTGCTAGGTTGCAGGCTGTAGTTTACAAAAACGAATATGTCAGTGAATATGTGATTCTTTAA
 GGCTGCAATACAAGCATTGAGTTCCTGTTTCAATAAGAGTCAATCCACATTTACAAAGATG
 CATTTTCTTTTGTATAAAAGCAAATAATATGCTTCAGATTATTCTCTCAAATA
 TAACACATATCTAGATTTTCTGCTTGCATGATATTCAGGTTTCAGGAATGAGCCTTGTAAAT
 ATAAGTGGCTGTGACGCTCTGCTTCTCTTCTGTAAGTTTCAGCATGGGTGTGCCTTCATAC
 AATAATATTTTCTCTTTTGTCTCCAATAATAATAAATGTTTGTCTAAATCTTACAATTTGA
 AAGTAAAAATAAACAGAGTGATCAAGTTAAACCATACTACTATCTCTAAGTAAACGAAGGAGC
 TATTGGACTGTAAAAATCTCTCTGCACTGACAAATGGGGTTTGAAGATTTTGGCCCCACACT
 AACTCAGTTCTGTGTGATGAGAGACAATTAATAACAGTATAGTAATAATACCATATGATTTC
 TTTAGTTGTAGCTAAATGTTAGATCCACCGTGGGAAATCATTCCCTTTAAATGACAGCACA
 GTCCACTCAAAGGATTGCTAGCAATACAGCATCTTTTCTTTCACATGTCGAAGCCAAAAA
 TTTTAAGATGATTTGTGAGAAAGGACAAAAGTCCATACCCCTAATATTACAAGATTTGGTA
 AGCGCTCATCATTAATTTTATTTGTGGCAGGTATATTGACAGTTCGACCTGGAGGGTATGGA
 TATGGATATGAGCGTTCCAGAGACTATAATGGCAGAAACAGGGTGGTTATGACCGCTACTC
 AGGAGGAAATTACAGAGACAATTATGACAACTGAAATGAGACATGCACATAATATAGATACA
 CAGGAAATAATTTCTGATCCAGGATCGTCTTCCAAATGGCTGTATTATAAAGTTTGTGG
 AGCTGCACTGAAGCATCTTATTTTATAGTATACACCTTTTGTGTTTAAATTTAACTGCTGCA
 AGGTAGCTGAAGACCTTTTAGACAGTTCATCTTTTTTTAAATTTTCTGCTTATTTAA
 AGACAAATTATGGGACGTTTGTCAAAAAAAAAAAAAAAAAAAAAAAAAA

1007191.102461

FIGURE 230

MMLLVQGACCSNQWLA AVL LSLCCLLP SCLPAGQSVDFPWA AVDNMMV RKGDTAVLR CYLED
GASKGAWLNRSSI IFAGGDKWSVDPRVSI STL NKR DYSLQIQ NV DVTDDG PYTCSVQTQHTP
RTMQVHLTVQVPPKIYDISNDMTVNEGTVNLTCLATGKPEPSISWRHISPSAKPFENGQYL
DIYGITRDQAGEYECSAENAVSFDPVRKVKVVVNFAPT IQEIKSGTVTPGRSGLIRCEGAGV
PPPAFEWYKGEKKLFNGQQGII IQNFSTRSILTVTNVTQEHFGNYTCVAANKLGTTNASLPL
NPPSTAQYGITGSADVL FSCWYLVLTLS SFTSIFYLKNAILQ

Important features of the protein:

Signal peptide:

amino acids 1-31

Transmembrane domain:

amino acids 326-345

N-glycosylation sites.

amino acids 71-75, 153-157, 273-277, 284-288, 292-296, 305-309

Casein kinase II phosphorylation site.

amino acids 147-151, 208-212, 224-228

Tyrosine kinase phosphorylation site.

amino acids 178-186

N-myristoylation sites.

amino acids 7-13, 63-70, 67-73, 151-157, 239-245, 291-297,
302-308, 319-325

Myelin P0 protein:

amino acids 92-121

FIGURE 231

AGTGGTTCGATGGGAAGGATCTTTCTCCAAGTGGTTCCTCTTGAGGGGAGCATTCTGCTGG
CTCCAGGACTTTGGCCATCTATAAAGCTTGGCAATGAGAAATAAGAAAATTCTCAAGGAGGA
CGAGCTCTTGAGTGAGACCCAAAGCTGCTTTTACCCAAATTGCAATGGAGCCTTTCGAAA
TCAATGTTCCAAAGCCCAAGAGGAGAAATGGGGTGAACTTCTCCCTAGCTGTGGTGGTCATC
TACCTGATCCTGCTCACCGCTGGCGCTGGGCTGCTGGTGCTCCAAGTTCTGAATCTGCAGGC
GCGGCTCCGGGTCTTGAGATGTATTTCTCAATGACACTCTGGCGGCTGAGGACAGCCCGT
CCTTCTCCTGTGTCAGTCAGCACACCCTGGAGAACCTGGCTCAGGGTGCATCGAGGCTG
CAAGTCTGCAGGCCCAACTCACCTGGGTCCGCGTCAGCCATGAGCACTTGCTGCAGCGGGT
AGACAACTTCACTCAGAACCCAGGGATGTTCAGAATCAAAGGTGAACAAGGCGCCCCAGGTC
TTCAAGGTCACAAGGGGGCCATGGGCATGCCTGGTGCCCCCTGGCCCCCGGGAACCACTGCT
GAGAAGGGAGCCAAAGGGGGCTATGGGACGAGATGGAGCAACAGGCCCCCTCGGAGCCCAAGG
CCCACCGGAGTCAAGGGAGAGGCGGGCCTCCAAGGACCCAGGGTGCTCCAGGGAAGCAAG
GAGCCACTGGCACCCAGGACCCCAAGGAGAGAAGGGCAGCAAAGGCGATGGGGGTCTCATT
GGCCCAAAAGGGGAAACTGGAATAAGGGAGAGAAAGGAGACCTGGGTCTCCAGGAAGCAA
AGGGGACAGGGGCATGAAAGGAGATGCAGGGGTTCATGGGGCTCCTGGAGCCAGGGGAGTA
AAGGTGACTTCGGGAGGCCAGGCCCCACAGGTTTGGCTGGTCTTCTGGAGCTAAAGGAGAT
CAAGGACAACTGGACTGCAGGGTGTTCCGGGCCCTCCTGGTGCACTGGGACACCCAGGTGC
CAAGGGTGAGCCTGGCAGTGCTGGCTCCCTGGGCGAGCAGGACTTCCAGGGAGCCCCGGGA
GTCAGGAGCCACAGGCCTGAAAGGAAGCAAAGGGGACACAGGACTTCAAGGACAGCAAGGA
AGAAAAGGAGAATCAGGAGTTCCAGGCCCTGCAGGTGTGAAGGGAGAACAGGGGAGCCAGG
GCTGGCAGGTCCCAAGGGAGCCCCCTGGACAAGCTGGCCAGAAGGGAGACCAGGGAGTGAAAG
GATCTTCTGGGGAGCAAGGAGTAAAGGGAGAAAAAGGTGAAAGAGGTGAAAACCTCAGTGTC
GTCAGGATTGTTCGGCAGTAGTAACCGAGGCCGGGCTGAAGTTTACTACAGTGGTACCTGGGG
GACAATTTGCGATGACGAGTGGCAAAATTCTGATGCCATTGTCTTCTGCCGCATGCTGGGTT
ACTCCAAGGAAGGGCCCTGTACAAAGTGGGAGCTGGCACTGGGCAGATCTGGCTGGATAAT
GTTCACTGTTCGGGACGAGAGTACCTGTGGAGCTGCACCAAGAATAGCTGGGGCCATCA
TGACTGCAGCCACGAGGAGGACGCAGGCGTGGAGTGCAGCGCTCTGACCCGGAAACCCCTTCA
CTTCTCTGCTCCCGAGGTGTCTCGGGCTCATATGTGGGAAGGCAGAGGATCTCTGAGGAGT
TCCCTGGGGACAACCTGAGCAGCCTCTGGAGAGGGGCCATTAATAAAGCTCAACATCATTTGA

FIGURE 232

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA68886

><subunit 1 of 1, 520 aa, 1 stop

><MW: 52658, pI: 9.16, NX(S/T): 3

MRNKKILKEDELLSETQQAAPHQIAMEPFEINVPKPKRRNGVNFSLAVVVIYLILLTAGAGL
LVVQVLNLQARLRVLEMYFLNDTLAAEDSPSFSLLQSAHPGEHLAQGASRLQVLQAQLTWVR
VSHEHLLQVRDNTQNPGMFRIKGEQGAAPGLQGHKGAMGMPGAPGPPGPAEKGAKGAMGRD
GATGPSGPQGPPPGVKGEAGLQGPQGAAPGKQCATGTPGPQGEKGSKGDGGLIGPKGETGTKGE
KGDGLPLPGSKGDRGMKGDAVGMGPPGAQGSKGDGFRPGPPGLAGFPKAGDQGPGLQGVPG
PPGAVGHPGAKGEPGSAGSPGRAGLPGSPGSPGATGLKGSKGDGTLQGGQGRKGESGVPGPA
GVKGEQGSPLAGFPKGAAPGQAGQKGDQGVKGSSGEQGVKGEKGERGENSVSVRIVGSSNRGR
AEVYYSGTWGTICDDEWQNSDAIVFCRMLGYSKGRALYKVGAGTGQIWLNDNVQCRGTESTLW
SCTKNSWGHHDSCSHEEDAGVECSV

Transmembrane domain:

amino acids 47-66 (type II)

N-glycosylation sites.

amino acids 43-47, 83-87, 136-140

Tyrosine kinase phosphorylation site.

amino acids 432-440

N-myristoylation sites.

amino acids 41-47, 178-184, 253-259, 274-280, 340-346, 346-352,
400-406, 441-447, 475-481, 490-496, 515-521

Amidation site.

amino acids 360-364

Leucine zipper pattern.

amino acids 56-78

Speract receptor repeat

amino acids 422-471, 488-519

Clq domain proteins.

amino acids 151-184, 301-334, 316-349

FIGURE 233

CCCACGCGTCCGAAGGCAGACAAAGGTTCAATTTGTAAGAAGCTCCTTCCAGCACCTCCTCT
CTTCTCCTTTTGCCCAAACCTACCCAGTGAGTGTGAGCATTTAAGAAGCATCCTCTGCCAAG
ACCAAAAGGAAAGAAGAAAAAGGGCCAAAAGCCAAATGAACCTGATGGTACTTGTTTTCAC
CATTGGGCTAACTTTGCTGCTAGGAGTTCAAGCCATGCCTGCAAATCGCCTCTCTTGCTACA
GAAAGATACTAAAAGATCACAACTGTCAACCTTCCGGAAGGAGTAGCTGACCTGACACAG
ATTGATGTCAATGTCCAGGATCATTTCTGGGATGGGAAGGATGTGAGATGATCTGTTACTG
CAACTTCAGCGAATTGCTCTGCTGCCCAAAGACGTTTCTTTGGACCAAAGATCTCTTTG
TGATTCCCTTGCAACAATCAATGAAGATCTTCATGTATTCTGGAGAACACCATTTCCTGATTTC
CCACAACTGCACTACATCAGTATAACTGCATTTCTAGTTTCTATATAGTGAATAGAGCAT
AGATTCTATAAATTCCTTACTTGTCTAAGACAAGTAAATCTGTGTTAAACAAGTAGTAATAAA
AGTTAATTCAATCTAAAAAAAAAAAAA

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

FIGURE 234

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52758
<subunit 1 of 1, 98 aa, 1 stop
<MW: 11081, pI: 6.68, NX(S/T): 1
MKLMVLVFTIGLTLLLGVDQAMPANRLSCYRKILKDHNCHNLPEGVADLTQIDVNVQDHFWDG
KGCEMICYC�FSELLCCPKDVFFGPKISFVIPCNNQ
```

Important features:

Signal peptide:

amino acids 1-20

N-glycosylation site.

amino acids 72-76

Tyrosine kinase phosphorylation site.

amino acids 63-71

FIGURE 235

CCCACGCGTCCGCGGACGCGTGGGCTGGACCCAGGTCTGGAGCGAATTCAGCCTGCAGG
 CTGATAAGCGAGGCATTAGTGAGATTGAGAGAGACTTACCCCGCGTGGTGGTTGGAGGGC
 GCGCAGTAGAGCAGCAGCACAGGCGCGGGTCCCGGGAGGCCGGCTCTGCTCGCGCCGAGATG
 TGGAAATCTCCTTCACGAAACCGACTCGGCTGTGGCCACCGCGCGCCCGCGCTGGCTGTG
 CGCTGGGGCGCTGGTGTGGCGGGTGGCTTCTTTCTCGGCTTCCTCTTCGGGGTGGTTTA
 TAAAAATCCTCCAATGAAGCTACTAACATTACTCTCAAAGCATAATATGAAGCATTATTTGGAT
 GAATTTGAAGCTGAGAACATCAAGAAGTTCTTACATAATTTTACACAGATACCACATTTAGC
 AGGAACAGAAACAAAACCTTTAGCTTGCAAAGCAAATTCATCCCAGTGGAAAGAATTTGGCC
 TGGATTCTGTTGAGCTAGCTCATTTATGATGTCCTGTTGTCTACCCAAATAAGACTCATCCC
 AACTACATCTCAATAATTAAATGAAGATGGAAATGAGATTTTCAACACATCATTTATTGAACC
 ACCTCCTCCAGGATATGAAAATGTTTCGGATATTTGTACCACCTTTTCAGTGCTTTCTCTCCTC
 AAGGAATGCCAGAGGGCGATCTAGTGTATGTTAACTATGCAAGCACTGAAGACTTCTTTAAA
 TTGGACCGGGACATGAAAATCAATTGCTCTGGGAAAATTTGTAATTGCCAGATATGGGAAAGT
 TTTCCAGAGGAAATAAGTTAAAAATGCCAGCTGGCAGGGGCCAAAGAGTCACTCTCTACT
 CCGACCTCTGACTACTTTTGTCTCTGGGGTGAAGTCTCTACAGACGGTTGGAATCTCTCTC
 GGAGTGTGTCCAGCGTGGAAATATCCTAAATCTGAATGGTGACGAGAGACCTCTCACACC
 AGGTTACCCAGCAAAATGAATATGCTTATAGGCGTGGAAATTCAGAGGGCTGTTGGTCTTCCAA
 GTATTCTCTGTTTCATCCAATTGGATACTATGATGCACAGAAGCTCTAGAAAAAATGGGTGGC
 TCAGCACCAACAGATAGCAGCTGGAGAGGAAGTCTCAAAGTGCCCTACAATGTTGGACCTGG
 CTTTACTGGAACCTTTTCTACACAAAAGTCAAGATGCACATCCACTCTACCAATGAAGTGA
 CGAGAAATTTACAAATGTGATAGGTACTCTCAGAGGAGCAGTGGAAACAGACAGATATGTCAAT
 CTGGGAGGTCAACGGGACTCATGGGTGTTTGGTGGTATTGACCCCTCAGAGTGGAGCAGCTGT
 TGTTCATGAAATTTGTGAGGAGCTTTGGAACACTGAAAAAGGAAGGGTGGAGACCTAGAAGAA
 CAATTTTGTGTTGCAAGCTGGGATGCAGAAGAATTTGGTCTTCTTGGTTCCTAGAGTGGGCA
 GAGGAGAATTCAGACTCCTTCAAGAGCGTGGCGTGGCTTATATTAATGCTGACTCATCTAT
 AGAAGGAACTACACTCTGAGAGTTGATTTGACACCGCTGATGTACAGCTTGGTACACAACC
 TAAACAAAAGAGCTGAAAAGCCCTGATGAAGGCTTTGAAGGCAAACTCTCTTTATGAAGAGTTGG
 ACTAAAAAAGTCCTTCCCCAGAGTTTCAAGTGGCATGCCAGGATAAGCAAATGGGATCTGG
 AAATGATTTTGAGGTGTTCTTCCAACGACTTGAATTTGCTTCAGGCAGAGCAGCGTATACTA
 AAAATTTGGGAAACAAACAAATTCAGCGCTATCCAATGTATCAAGTGTCTATGAACATAT
 GAGTTGGTGGAAAAGTTTATGATCCAATGTTTAAATATCACCTCACTGTGGCCAGGTTTCG
 AGGAGGGATGGTGTGTTGAGCTAGCCAATCCATAGTGCTCCCTTTGATTGTGAGATTATG
 CTGTAGTTTAAAGAAATGCTGACAAAATCTACAGTATTTCTATGAACATCCACAGGAA
 ATGAAGACATACAGTGATCATTTGATTCACTTTTCTGCAGTAAAGAAATTTACAGAAAT
 TGCTTCCAAGTTTCAGTGAGAGACTCCAGGACTTTGACAAAAGCAACCCAAATGATATTGAAG
 TGATGAATCAACTCATGTGTTTCTGGAAGAGCAATTTATGATCAATTAGGTTACAGAC
 AGGCCTTTTATAGGCATGTCTATGCTCCAAGCAGCCACAACAAGTATGCAGGGGAGTC
 ATTCCAGGAATTTATGATGCTCTGTTTGATATTGAAGCAAAAGTGGACCTTCCAAGGCTC
 GGGGAGAAGTGAGAGACAGATTTATGTTGCAGCCTTCACAGTGAGGAGCTGCAGAGACT
 TTGAGTGAAGTAGCCTAAGAGGATTTTATGAGAATCCGATTGAATTTGTGTGGTATGTCA
 CTCAGAAAGAATCGTAATGGGTATATTGATAAAATTTAAAAATGGGTATATTGAAATAAAGT
 TGAATATTATATATAA

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000

FIGURE 236

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA52756

><subunit 1 of 1, 750 aa, 1 stop

><MW: 84305, pI: 6.93, NX(S/T): 10

MWNLLHETDSAVATARPRWLCAGALVLGGFFLLGFLFGWFIKSSNEATNITPKHNMKAFL
DELKAENIKKFLHNFTQIPHLAGTEQNFLAKQIQSQWKEFGLDSVELAHYDVLLSYPNKTH
PNYISIIINEDGNEIFNTSLFEPPPPPGYENVSDIVPPPSAFSPQGMPEGDLVYVNYARTEDFF
KLERDMKINC SGKIV IARYGKVFRGNKVNAQLAGAKGVILYSDPADYFAPGVKSYPDGWNL
PGGGVQVRGNILNLNGAGDPLTPGYPAN EYAYRRGIAEAVGLPSIPVHPIGYYDAQKLEKMG
GSAPPDSSWRGSLKVPYNVGPFGFTGNFSTQKVKMHIHSTNEVTRIYNVIGTLRGAVEPD RYV
ILGGHRDSWVFGGIDPQSGAAVVHEIVRSFGTLKKEGWRPRRTILFASWDAEEFGLLGSTEW
AEENSRLQLQERG VAYINADSSIEGNYTLRVDCTPLMYSLVHNLT KELKSPDEGFEGKSLYES
WTKKSPSP EFGMPRISKLGSGNDFEVFFQRLGIASGRARYTKNWETNKFSGYPLYHSVYET
YELVEKFYDPMFKYHLTVAQVRGGMVFELANSIVLPFDCCR DYAVVLRKYADKIYSISMKH PQ
EMKTYSVSFDLSFSAVKNFTEIASKF SERLQDFDKSNPIVLRMMNDQLMFLERAFIDPLGLP
DRPFYRHVIYAPSSHNKYAGESFPGIYDALFDIESKVDPSKAWGEVKKRIYVAAFTVQAAAE
TLSEVA

Signal sequence:

amino acids 1-40

N-glycosylation sites.

amino acids 76-80, 121-125, 140-144, 153-157, 195-199, 336-340,
459-463, 476-480, 638-642

Tyrosine kinase phosphorylation sites.

amino acids 363-372, 605-613, 606-613, 617-626

N-myristoylation sites.

amino acids 85-91, 168-174, 252-258, 256-262, 282-288, 335-341,
360-366, 427-433, 529-535, 707-713